Biochemical characterization of the relaxase TraI, the coupling protein TraD and the hypothetical protein Yaf of the novel Type IV Secretion System from the human pathogen *Neisseria gonorrhoeae*

Philipps Universität Marburg

DISSERTATION

zur Erlangung des Doktorgrades der Naturwissenschaften

(Dr. rer. nat.)

Dem Fachbereich Biologie

der Philipps-Universität Marburg

vorgelegt von

Eva-Maria Heller

aus Schwalmstadt

Marburg an der Lahn, November 2012

Die Untersuchungen zur vorliegenden Arbeit wurden von Dezember 2008 bis September 2012 am Max-Planck-Institut für terrestrische Mikrobiologie in Marburg an der Lahn in der Abteilung Ökophysiologie unter der Leitung von Dr. Christiaan van der Does durchgeführt.

Im Fachbereich Biologie der Philipps-Universität Marburg als

Dissertation angenommen am: 04.12.2012

Erstgutachter: Prof. Dr. Lotte Søgaard-Andersen

Zweitgutachter: Prof. Dr. Erhard Bremer

Weitere Mitglieder der Prüfungskommision: Prof. Dr. Uwe Maier

Prof. Dr. Michael Bölker

Tag der mündlichen Prüfung: 05.02.1013

Hiermit versichere ich, dass die vorliegende Dissertation mit dem Titel

"Biochemical characterization of the relaxase TraI, the coupling protein TraD and the hypothetical protein Yaf of the novel Type IV Secretion System from the human pathogen *Neisseria gonorrhoeae*"

selbstständig verfasst, keine anderen als die im Text angegeben Hilfmittel verwendet und sämtliche Stellen, die im Wortlaut oder dem Sinn nach aus anderen Werken entnommen sind, mit Quellenangaben kenntlich gemacht habe.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den 02.11.2012 Eva-Maria Heller

Table of Content

I. Summary

Type IV secretion systems (T4SSs) are able to transport DNA and effector proteins across the inner and outer membranes of prokaryotes. T4SSs consist of a transport complex which forms the channel across the inner and outer membrane, the so called mating pair formation complex (MPF), and DNA processing and transfer proteins (Dtr proteins). A unique T4SS is encoded in the gonococcal genomic island which is found in most *Neisseria gonorrhoeae* strains. Remarkably, this secretion system is the first T4SSs that was shown to secrete ssDNA during all phases of growth directly into the environment. This property might facilitate functional studies on the transport mechanisms used by T4SSs. Remarkably, the T4SS encoded within the GGI consists of a MPF complex that is homolgous to the MPF complex of the *E. coli* F-plasmid (MPF_F family) and Dtr proteins that belong to the MOB_H family of targeting factors. Although many different members of the MOB_H family have already been identified, almost no information is currently available on this novel group of targeting factors.

To study the conservation of GGI-like T4SSs, similar T4SS were identified in a computational analysis of all currently completely sequenced chromosomes and plasmids. Next to *N. gonorrhoeae* very similar conserved T4SSs have been found in several other bacteria. These T4SSs are either encoded within the chromosome or on plasmids. Most of the genes encoded in the GGI between the *ltgX* and *exp1* genes, the region which encodes mainly the MPF_F proteins, were highly conserved in allmost all of the identified systems, even though, several of theses genes have been shown previously not to be important for ssDNA transport. Remarkably, all identified chromosomal GGI-like T4SSs encode *mpf* genes of the MPF_F family and *dtr* genes of the MOB_H family; whereas, the plasmid encoded GGI-like T4SSs always contain *dtr* genes of the MOB_F family and *mpf* genes of the MPF_F family. This suggests that relaxases of the MOB_H family are preferentially found in chromosomally encoded T4SSs.

To elucidate the functional mechanism of the DNA processing reaction of this unique DNA secretion system a biochemical approach was set up to functionally characterize the coupling protein TraD, the relaxase TraI and the hypothetical protein Yaf which are different key enzymes of the Dtr proteins of *N. gonorrhoeae*.

The overexpression and purification of coupling proteins and relaxases of other T4SSs has previously been shown to be inherently difficult. Although many different approaches using *E. coli* as host for a recombinant overexpression have been tested, it was not possible to obtain soluble TraD and large amounts of inclusion bodies were produced. Small amounts of TraD have been detected in isolated inner membranes of *E. coli*, but all performed solubilization attempts failed. Finally, TraD was isolated and purified from inclusion bodies under denaturing conditions, but all further refolding approaches with the purified protein resulted either in the loss or the aggregation of TraD. Therefore, it was not possible to characterize the enzymatic activity of the neisserial coupling protein TraD.

The hypothetical protein Yaf, which is encoded in same operon with *traI* and *traD* was sucessfully overexpressed and purified. Interestingly, Yaf forms different oligomeric states in solution. The most abundant oligomer of native Yaf was the dimeric state. The dynamic equilibrium of the different oligomeric states was strongly dependent on the concentration of the protein and the buffer conditions. The interconversion between the different oligomeric states is a slow process. The presence of an N-terminal His-tag strongly stabilized the tetrameric form. Limited proteolysis experiments showed that Yaf is highly resistant against trypsin and most likely has a very compact

and stable structure. Tetrameric Yaf showed a higher resistance against proteolytic digestion than dimeric Yaf. Remarkably, the introduction of different point mutations into the sequence of Yaf resulted in a destabilization of the protein and a strong decrease in the amount of soluble protein. Different activity assays have been performed with the purified protein to study the function of Yaf. DNA binding assays showed binding of dimeric and tetrameric Yaf to dsDNA with low affinity (>5 µM). Furthermore, a metal dependent sequence unspecific nuclease activity was observed associated with elution fractions containing dimeric and tetrameric Yaf. This nuclease activity was specific for dsDNA fragments and oligonucleotides that contained secondary structures. In the presence of different divalent metals, a relaxation activity on supercoiled plasmid DNA has been observed for fractions containing dimeric and tetrameric Yaf, and was strongly enhanced in the presence of Mn^{2+} and Co^{2+}, but was also observed in the presence of Mg^{2+}. To obtain further information on the structure and function of Yaf, crystallization attempts have been started in close collaboration with Dr. S. Smits in the group of Prof. Dr. L. Schmitt (Heinrich Heine University in Düsseldorf). Small crystals of N-terminally His-tagged Yaf diffracted up to 3 Å. Remarkably, no crystals have been obtained for native Yaf in any of the tested conditions. Recentyl, crystals of N-terminal His-tagged Yaf with incorporated selenomethionine have been obtained.

To characterize the MOB_H relaxase TraI, many different overexpression and purification approaches have been tested. Finally, an overexpression and purification method could be established to purify soluble TraI. Isolated TraI formed stable dimers in solution, and limited proteolysis revealed the presence of at least three trypsin resistant domains in TraI. DNA binding assays showed that TraI binds to ssDNA and dsDNA with high affinity in a sequence unspecific manner. The minimal binding frame of TraI for ssDNA substrates has been determined with 9 nucleotides. Furthermore, a metal dependent sequence unspecific relaxation activity of TraI on supercoiled plasmids has been observed, with Mn^{2+} as preferred metal cofactor. The relaxation activity of TraI is similar to the activity of *E. coli* Topoisomerase I, but differs within the preferred metal cofactor and the lack of detectable intermediates during the relaxation reaction of TraI. DNA cleavage assays performed with different ssDNA substrates demonstrated that TraI cleaves oligonucleotides containing the putative *oritT* sequence of the GGI or the *oriT* region of the *E. coli* F-plasmid strand specific in a Mn^{2+} dependent manner. The complementary strands and a polyT oligonucleotide were not cleaved by TraI. Analysis of the putative cleavage sites did not show any sequence similarity between both cleaved oligonucleotides. To obtain structural information about TraI, crystallization attempts in close collaboration with Dr. S. Smits in the group of Prof. L. Schmitt (Heinrich Heine University in Düsseldorf) were started.

Besides the characterization of the enzymatic activity of TraI and Yaf, pull down assays were performed to test whether these proteins interact with each other or with the T4SS ATPase TraC. Unfortunately, no protein-protein interactions were detected between these three proteins.

The experimental results from the current study help to understand the DNA processing mechanism of the unique T4SS of *N. gonorrhoeae*. Important steps have been made in the functional overexpression and the purification of Yaf and TraI. The hereby obtained results from the biochemical characterization of TraI represent the first reported biochemical data which is available on the novel MOB_H relaxase family. Furthermore, this study revealed first insights into a possible enzymatic activity of the previously uncharacterized hypothetical protein Yaf, even though the function of this protein still remains unknown.

II. Zusammenfassung

Typ IV Sekretionssysteme (T4SSs) transportieren DNA und verschiedenen Effektorproteine über die innere und äußere Zellmembran von Prokarioten. Sie bestehen aus einem Transportkomplex, dem sogenannten MPF-Komplex („mating pair formation“), der einen Kanal durch die innere und äußere Membran bildet und aus DNA Prozessierungs- und Transferproteine (Dtr Proteine).

Ein neues und bisher einzigartiges T4SS ist innerhalb des „gonococcal genomic island“ lokalisiert, welchen in den meisten *Neisseria gonorrhoeae* Stämmen vorkommt. Dieses T4SS ist das erste T4SS, das einzelsträngige DNA in allen Wachstumsphasen direkt in seine Umgebung sekretiert. Diese Eigenschaft macht das T4SS von *N. gonorrhoeae* zu einen vielversprechendem Forschungsobjekt für Studien über den Transportmechanismus von T4SSs. Das GGI-T4SS besteht aus einem MPF-Komplex, der große Ähnlichkeit mit dem MPF-Komplex des F-Plasmides von *E. coli* hat (die sogenannte MPF_F Familie), und aus Dtr Proteinen die zu der MOB_H Familie der Prozessierungs- und Lokalisationsfaktoren gehören.

Um die Verteilung und Konservierung GGI ähnlicher T4SSs zu untersuchen wurden mittels bioinformatischer Analysen ähnliche T4SSs aus allen derzeitig sequenzierten Genomen und Plasmiden identifiert. Neben dem T4SS aus *N. gonorrhoeae* wurde weitere ähnlich konservierte T4SSs in anderen Bakterien identifiziert. Diese T4SSs sind entweder genomisch oder befinden sich auf Plasmiden. Die meisten der GGI Gene innerhalb der Region zwischen *ltgX* und *exp1* kodieren für MPF_F Proteine und sind hoch konserviert in allen identifizierten T4SSs, auch wenn einige dieser Gene nicht in den Transfer von Einzelstrand-DNA involviert sind. Bemerkenswerterweise enthalten alle identifizierten genomischen T4SSs *mpf* Gene der MPF_F Familie und *dtr* Gene der MOB_H Familie; die Plasmid-kodierten T4SSs hingegen enthalten *mpf* Gene der MPF_F Familie und *dtr* Gene der MOB_F Familie. Die zeigt, dass Relaxasen der MOB_H Familie hauptsächlich in genomischen T4SSs vorhanden sind.

Um den Mechanismus der DNA Prozessierung des einzigartigen T4SSs aus *N. gonorrhoeae* zu studieren wurden in dieser Arbeit die Schlüsselenzyme TraD (das Kopplungsprotein), TraI (die Relaxase) und das hypothetische Protein Yaf aus *N. gonorrhoeae* biochemisch charakterisiert.

Die Überexpression und die Aufreinigung von Kopplungsproteinen und Relaxasen von T4SSs ist eine schwierige Herausforderung und obwohl viele verschiedene Versuche unternommen wurden um TraD rekombinant in *E. coli* zu überexprimieren, war es nicht möglich lösliches Protein zu isolieren. Stattdessen wurden große Mengen von Einschlusskörperchen gebildet. Kleine Mengen von TraD wurden in isolierten inneren Zellmembranen von *E. coli* detektiert, allerdings konnte dieses Protein nicht aus den Membranen herausgelöst werden. Schließlich wurde TraD unter dentaurierenden Bedingungen aus Einschlusskörperchen isoliert und aufgereinigt, allerdings schlugen alle weiteren Neufaltungsversuche fehl und führten entweder zum Verlust des Proteins oder zu Aggregation. Deshalb war es nicht möglich die enzymatische Aktivität des neisserialen Kopplunsproteins zu charakterisieren.

Das hypothetische Protein Yaf konnte erfolgreich in großen Mengen isoliert werden und bildet in Lösung ein Gleichgewicht zwischen verschiedenen oligomeren Stadien. Hauptsächlich werden jedoch Dimere gebildet. Das Gleichgewicht zwischen den verschiedenen Oligomeren wird deutlich von der Proteinkonzentration und den jeweiligen Pufferbedingungen beeinflusst. Die Umwandlung der

einzelnen Oligomerstadien ist ein langsam verlaufender Prozess. Die Anwesenheit eines N-terminalen His-tags hat einen stabilisierenden Effekt auf die tetramere Form. Limitierte Proteolyse mit Trypsin zeigte eine hohe Resistenz von Yaf gegenüber Trypsin, was auf eine hohe strukturelle Stabilität von Yaf hindeutet. Tetramere zeigten hierbei eine größere Resistenz gegenüber der proteolytischen Aktivität von Trypsin als Dimere. Erstaunlicherweise führten Punktmutationen innerhalb der Sequenze von Yaf zu einer Destabilisierung des Proteins und die Mengen an löslichem Protein waren stark reduziert. Mit Hilfer verschiedener Aktivitätsmessungen wurde die Funktion von Yaf untersucht. DNA-Bindungsversuche haben gezeigt, dass dimeres und tetrameres Yaf doppelsträngige DNA Substrate mit geringer Affinität (>5 µM) bindet. Desweiteren wurde in Elutionsfraktionen mit dimeren und tetrameren Yaf eine Metall-abhängige und Sequenz-unabhängige Nukleaseaktivität beobachtet. Diese Nukleaseaktivität war spezifisch für doppelsträngige DNA Fragmente und für Oligonukleotide mit Sekundärstrukturen. In der Anwesenheit von verschiedenen divalenten Metallionen konnte innerhalb der Elutionsfraktionen mit dimerem und tetramerem Yaf eine Relaxierungsaktivität auf superspiralisierte Plasmide beobachtet werden. Diese Reaktion war besonders verstärkt in Anwesenheit von Mn^{2+} und Co^{2+}, konnte aber auch in Anwesenheit von Mg^{2+} beobachtet werden. Um Informationen über den strukturellen Aufbau und über die Funktion von Yaf zu erhalten, wurde die Kristallisation von Yaf in Zusammenarbeit mit Dr. S. Smits aus der Forschungsgruppe von Prof. L. Schmitt (Heinrich Heine Universität in Düsseldorf) begonnen. Kristalle, die aus Yaf mit einem N-terminalen His-tag bestehen, streuen zurzeit bis zu einer Auflösung von 3 Å. Zudem konnte kürzlich erstes Wachstum von Protein-Kristallen mit inkorporiertem Selenomethionin beobachtet werden. Natives Yaf kristallisiert jedoch nicht unter den gleichen Bedingungen wie Yaf mit einem N-terminalen His-tag.

Um die MOB_H Relaxase TraI funktionell zu charakterisieren wurden viele verschiedene Überexpressionsmethoden und Reinigungsprotokolle getestet. Schließlich konnten eine Überexpressionsmethode und ein Reinigungsprotokoll etabliert werden, mit denen erfolgreich lösliches TraI isoliert werden konnte. TraI bildet in Lösung stabile Dimere und limitierte Proteolyse zeigte, dass TraI aus mindestens drei unterschiedlichen Trypsin-resistenten Domänen besteht. DNA-Bindgungsversuche haben gezeit, dass TraI doppelsträngige und einzeilsträngige DNA mit hoher Affinität in einer Sequenz-unabhängigen Reaktion bindet. Die minimale Bindungslänge von einzelsträngigen DNA Substraten, die für eine Bindung mit TraI nötig ist wurde auf 9 Nukeotide bestimmt. Desweiteren wurde für TraI eine Metall-abhängige und Sequenz-unabhängige Relaxierungsaktivität auf superspiralisierte Plasmide beobachtet. In Anwesenheit von Mn^{2+} ist diese Aktivität deutlich gesteigert. Die Relaxierungsaktivität von TraI ist ähnlich zu der enzymatischen Aktivität von Topoisomerase I aus *E. coli*, unterscheidet sich aber von dieser bezüglich des bevorzugten Metall-Cofaktors und durch die tatsache, dass bei der Reaktion von TraI keine detektierbaren Topoisomerasen I-typischen Intermediate während der Reaktion beobachtet wurden. Experimente zur DNA-Strangspaltungsaktivität von TraI wurden mit unterschiedlichen einzel-strängigen DNA Substraten durchgeführt. Dabei wurde ein TraI induzierte Strang-spezifische Spaltung bei zwei Oligonukleotiden beobachtet, die abhängig ist von Mn^{2+}. TraI spaltet einzelsträngige DNA Substrate, die entweder die Sequenze der *oriT* Region des GGI enthalten oder die Sequenz der *oriT* Region des F-Plasmides aus *E. coli*. Dazu komplementäre Stränge und ein polyT Oligonukleotid wurden nicht von TraI gespalten. Ein Vergleich der beiden postulierten Spaltstellen in den Oligonukleotiden ergab keine Ähnlichkeiten zwischen den beiden Sequenzen. Um strukturelle Informationen über TraI zu erhalten, wurden die Kristallisation von TraI in Zusammenarbeit mit

Dr. S. Smits aus der Forschungsgruppe von Prof. L. Schmitt (Heinrich Heine Universität in Düsseldorf) initialisiert.

Neben der Charakterisierung der enzymatischen Aktivität von TraI und Yaf wurde mit „pull-down“ Versuchen untersucht, ob beide Proteine miteinander oder ob sie mit mit der neisseriellen T4SS ATPase TraC interagieren. Die Ergebnisse zeigten jedoch keine Interaktion zwischen diesen drei Proteinen.

Die Ergebnisse dieser Studie führen dazu, ein besseres Verständnis bezüglich des DNA-Prozessierungsmechanismus des einzigartigen T4SS von *N. gonorrhoeae* zu erhalten. Bedeutende Fortschritte wurden bei der funktionellen Überexpression und der Aufreinigung von TraI und Yaf erzielt. Die, in dieser Studie erzielten Ergebnisse der biochemischen Charakterisierung von TraI repräsentieren die ersten verfügbaren biochemischen Daten über diese neue Relaxase-Familie. Desweitern ermöglicht diese Studie erstmals Einblicke in die enzymatische Aktivität des nicht charakterisierten hypothetischen Proteins Yaf, obwohl die eigentliche Funktion dieses Proteins noch immer unbekannt ist.

1. Introduction

1.1 *Neisseria gonorrhoeae*

The genus *Neisseria* belongs to the family of *Neisseriaceae* which are classified into the group of *Neisseriales* and belong to the class of Gram-negative β-proteobacteria. These are mostly strict aerobic coccal or diplococcal bacteria (see Figure 1). Most members of this family are commensal bacteria colonizing human mucosal tissues. Two species within this genus, *Neisseria gonorrhoeae* and *Neisseria meningitidis* are serious human pathogens causing diseases like gonorrhea, bacterial meningitis, or septicemia. Both are non flagellated diplococci and are either termed gonococcus (*N. gonorrhoeae*) or meningococcus (*N. meningitides*) (1-4). Albert Neisser described *N. gonorrhoeae* as the causative agent for gonorrhea first in 1879.

Gonorrhea is due to its rapid spread and serious symptoms, as well as often accompanying diseases (e.g. HIV), a very dangerous infection that fortunately still can be treated with antibiotics. In cases of co-infection with gonorrhea and HIV, *N. gonorrhoeae* is under suspicion to enhance the infection rate with HIV to the sexual partners and is therefore a potent amplifier of the spread of HIV (5). Globally more than 60 million new cases are recorded each year (WHO – World Health Organization), mostly appearing in less developed countries (6). High infection rates can be found within homosexually active men and women, as well as within indigenous populations and in poor areas (7).

The infection of human urogenital tissues with *N. gonorrhoeae* causes urethritis in men and endocervicitis in women. However, *N. gonorrhoeae* is also able to infect other mucosal tissues, mostly resulting in more difficult diagnosis and treatments of the diseases (e.g. anorectal or pharyngeal infections). One complicating aspect in diagnosing gonorrhea is the fact that most endocervical, anorectal and pharyngeal infection are asymptomatic and therefore difficult to realize. In case of an infection spreading to adjacent tissues, men can develop an epididymoorchitis, and in women it can result in pelivic inflammatory disease (PID) or in few cases in gonococcemia which could lead to tenosynovitis, septic arthritis, endocarditis, or meningitis. Within women even more dangerous complication are possible since a salpingitis can cause PID and infertility, or can lead to an infection of a newborn during delivery. Gonorrhea is also accountable for first-trimester abortion and complications in ectopic pregnancies (8).

The effective treatment of gonorrhea with antibiotics has become increasingly difficult since antibiotic resistances are mounting fast. Classical antibiotics like penicillins, tetracyclines, macrolides, spectinomycine, and even quinolones are already ineffective for use. In many industrial countries the recommended antibiotics for treatment of gonorrhea are therefore third-generation cephalosporins and ceftriaxone (8). The antibiotic resistance pattern within gonorrhea varies strongly in different countries, e.g. west pacific regions show resistance to penicillin, tetracyclines and quinolones, whereas in the Australian outback many strains are not resistant to penicillin (WHO). The spread of antibiotic resistances is caused by the natural competence of *N. gonorrhoeae* during all phases of growth, leading to a very efficient intra- and interspecies DNA transfer (9). In principle, DNA can be bound unspecific to the neisserial cell surface, but for effective DNA uptake and recombination the DNA needs to have a 10 bp conserved DUS sequence (DNA uptake sequence) (10). DUS sequences are accumulated in certain chromosomal areas, like e.g. regions encoding genes responsible for DNA maintenance and a few regions encoding other essential genes, implicating a potential function in

chromosome stability (11, 12). Besides this, it is in discussion whether or not recombination is a DNA repair mechanism in *Neisseria* since many typical DNA repair genes are lacking in this species (13). A comparison of genomes of three different *Neisseria* species revealed a clear differentiation between the species leading to the conclusion that interspecies recombination is an infrequent event (4). The DNA for transformation is released by autolysis, T4SS mediated secretion or blebbing (10, 14). The extraordinary T4SS of *N. gonorrhoeae* is located in the gonococcal genetic island (GGI) and will be described in more detail in Chapter 1.6. Blebbing occurs very frequently in *Neisseria* and seems to be an important mechanism to defend against the human immune system and to spread genetic information (15, 16). Blebs are detached membrane vesicles containing immunodominant cell surface antigens, as well as associated DNA and other attached proteins (14, 17). Dorward *et al.* showed the association of plasmid DNA to neisserial blebs and proposed the possibility of a new mechanism of interspecific plasmid transfer comparing the neisserial blebs with the transformasomes of *Haemophilus* (14, 18, 19). A large proportion of transformable DNA is supplied by the high incidence of autolysis within *Neisseria*. The mechanism of autolysis in *Neisseria* is up to now only poorly understood. The increasing antibiotic resistance and the resulting danger of incurability of gonorrhea make the understanding of the mechanism of intra- and interspecific DNA transfer an outstanding task. Recently, the *N. gonorrhoeae* strain H041 was listed as the first untreatable 'Superbug' of gonorrhoeae, since it is resistant to the commonly used antibiotics and even to Ceftriaxone (20, 21).

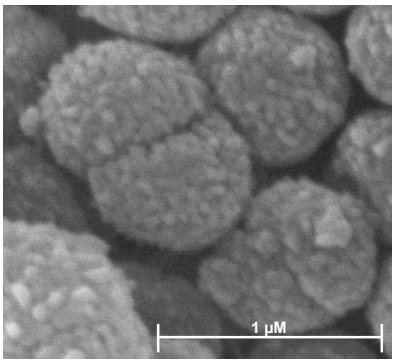


Figure 1: SEM picture of *N. gonorrhoeae* wild type strain MS11 with 20,000x magnitude. Cells were chemically fixed with glutaraldehyde and osmium tetroxide. The picture shows the typical diplococcal shape of *N. gonorrhoeae*. The micrograph was taken by T. Bender and A. Wagner (our group) with a Hitachi S-530 Scanning Electron Microscope with kind support of the group of Prof. Kost.

1.2 Bacterial secretion systems

Prokaryotes use secretion systems to secrete a broad range of substrates and have developed a huge set of different secretion systems. Up to now seven different secretion systems (type I-VII) are identified in Gram-negative and Gram-positive bacteria, as well as in archaea (22). Secretion systems are involved in versatile functions like e.g. intercellular signaling (23), bacterial movement (24, 25), effector protein transfer (26), symbiosis (27), regulation of metabolic pathways (27, 28), or DNA transfer (29, 30). Some of the systems are cell-contact dependent and transfer DNA and proteins into recipient cells (T3SS, T4SS, T5SS, T6SS), whereas other systems are not dependent on cell-contact and secrete surface molecules or effector proteins to the outer cell-surface and into the environment (T1SS, T2SS) (22, 31). Since T7SSs are the latest identified secretion systems, it is currently not known if their secretion mechanism is cell-contact dependent or not. So far T7SSs have only been identified in *Mycobacteria* and a few other Gram-negative bacteria and seem to be unique for this group (32), whereas other secretion systems (type I-VI) are widely spread among bacteria and archaea. A general

overview on the basic composition and organization of the prokaryotic secretion systems is given in Figure 2.

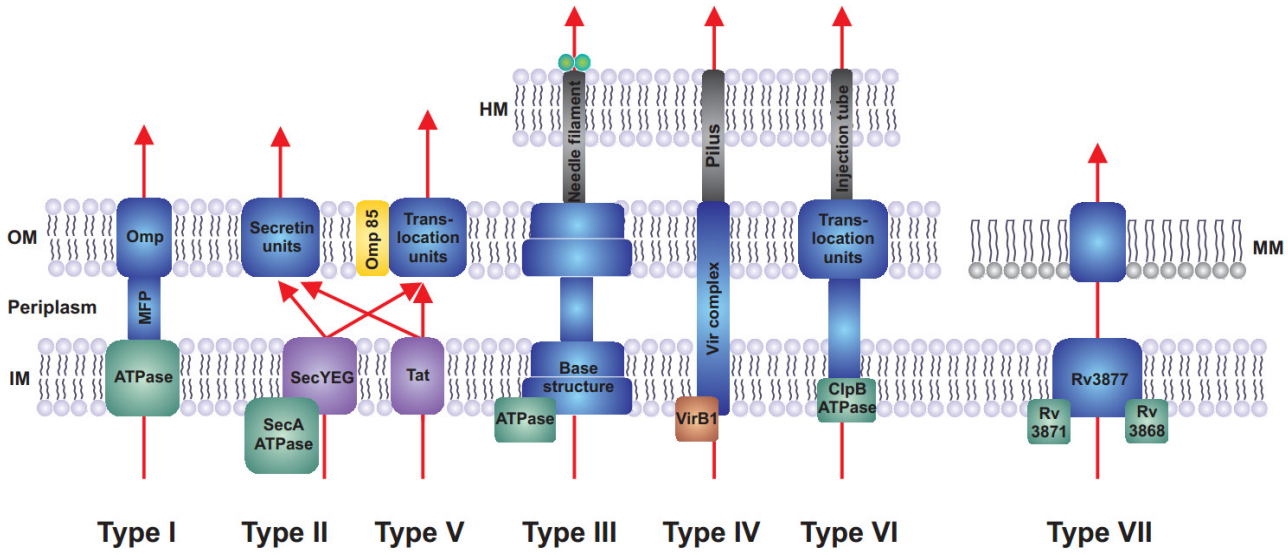


Figure 2: General overview of current models of the basic organization of bacterial secretion systems. Omp: outer membrane protein; MFP: membrane fusionprotein; OM: outer membrane; IM: inner membrane; MM: mycobacterial membrane; HM: host membrane. The scheme is adapted from (22).

Often several types of secretion systems are encoded within the genome of a species, e.g. *Pseudomonas aeruginosa* encodes five different secretion systems that are used for a broad set of functions (33). Each secretion system can also vary in its abundance and therefore several secretion systems of one type can be used for different functions within one organism. The secretion mechanisms of different secretion systems are either 1-step mechanisms (T1SS, T3SS, T4SS, and T6SS) in which the substrate is secreted in one step across the cell-envelope to the environment or into the host cell, or 2-step mechanisms (T2SS and T5SS) using the Sec- or Tat-translocation pathway prior secretion to deliver substrates into the periplasm followed by their translocation through the outer membrane via the secretion system (22, 31, 34).

The Sec-translocation pathway is commonly used by bacteria to transfer proteins across the cytoplasmic membrane, either into the periplasm in Gram-negative bacteria or to the cell-surface and environment in Gram-positive bacteria and archaea (22, 35). The translocation system consists of an inner-membrane-pore made of the three proteins SecY, SecE and SecG, as well as a cytoplasmic ATPase (SecA) which energizes besides the proton motive force (PMF) the translocation of unfolded proteins across the cytoplasmic membrane (36-38). The proteins are triggered to the translocation pore via an N-terminal signal sequence and translocation occurs either co-translationally (for membrane proteins or proteins with a very hydrophobic signal sequences) with the translating ribosome targeted and localized to the translocation pore, or the SecB chaperon targets the proteins after translation to the SecYEG pore via binding to SecA (22, 36, 38).

The Tat-translocation pathway is very similar to the SecYEG system. Bacteria and archaea use the Tat-system (twin-arginine-transporter) as an alternative translocation system to transfer folded proteins across the cytoplasmic membrane, contrary to the secretion of unfolded proteins in the SecYEG system (22, 39, 40). The Tat-system consists of three essential membrane proteins

(TatA, TatB and TatC). The translocation pore is made of the major pore-forming subunit TatA and PMF energizes the translocation process (39). Proteins for translocation contain a typical N-terminal recognition sequence (twin-arginine-motif: S-R-R-x-F-L-K) and are bound to a co-factor. They are recognized and targeted to the translocation-pore via TatC (39, 41, 42).

Type I secretion systems (T1SS) can be found in Gram-negative and Gram-positive bacteria and are relatively simple secretion systems with regard to their architecture and their secretion mechanism (43, 44). The substrates of T1SSs are often virulence factors like toxins or could be exopolysaccharides which are important for biofilm formation (45, 46). The secretion of the substrates into the environment is an ATP dependent mechanism in which the substrate is targeted to the secretion system via a C-terminal secretion signal that cannot be cleaved from the protein (47, 48). T1SSs consist of three different components which are an ABC-transporter (ATP binding cassette), an outer membrane factor, and a membrane fusionprotein that connects both membrane components (48). During T1S the proteins are secreted in an inactive unfolded state and are activated by changing their folding state at the outside of the cell (49). Some examples for T1SSs are the TolC-HlyB-HlyD system of *E. coli* (45), the Apr- and the Has-systems in *P. aeruginosa* (50, 51), or plant-pathogens like *Xanthomonas oryzae* pv. *oryzae* (52), *Agrobacterium tumefaciens* (53), *Pseudomonas syringae* pv. *tomato* (43), and *Ralstonia solanacearum* (43).

Type II secretion systems (T2SS), also known as secretons, are used by Gram-negative proteobacteria and have highly adopted secretion functions specific to the abiotic and biotic environment of the cell (33, 54). Often T2SSs are involved in the secretion of toxins and other virulence factors (22). These systems are unique in their ability to secrete large multimeric proteins in an ATP and Sec-translocation dependent 2-step-mechanism, which involves at least 11-16 different proteins (55-57). After the transfer via SecYEG the substrates are folded in the periplasm prior to secretion through the outer membrane. T2SSs are organized in three components with an integral inner membrane protein acting as a pore-anchor, a large outer membrane spanning channel formed of secretin subunits, and a type IV pilus like pseudopilus which has been proposed to act as a piston that pushes the substrates across the outer membrane (54-57). A cytosolic ATPase regulates the opening and closing of the pore. T2SSs can be found in many Gram-negative pathogens like e.g. the Xcp- and Hxc-systems of *P. aeruginosa* (58, 59), the StcE metalloprotease secretion system and the ADP-ribosylating toxin secretion system of *E. coli* (54), the cholera toxin of *Vibrio cholera* (54), or different substrates in *R. solanacearum*, *Legionella pneumophila*, *Klebsiella oxytoca*, or *Aeromonas hydrophila* (54).

The injectisome of type III secretion systems (T3SS) is evolutionary related to the bacterial flagellum and is used by Gram-negative pathogens or mutualists to deliver effector proteins into host cells (31, 60, 61). T3SSs can be separated in seven subfamilies and are often encoded in pathogenic islands or plasmids (60). They are large multiprotein systems with approximately 25 proteins involved in assembly and secretion. 9 of these 25 proteins are conserved in all subfamilies and 8 of these 9 proteins are also conserved in bacterial flagella systems (60, 62). T3S is a SecYEG independent 1-step-mechanism comparable to the injection with a syringe (60). The secretion system consists of an inner membrane embedded base structure with an inner and an outer ring and a cup-like structure at the cytosolic side, a needle-like filament mediating the contact with the host cell and the injection of virulence factors, and a pore which is assembled in the host membrane after cell-cell contact (31, 60, 63). The length of the needle can vary between different species and is very important for virulence,

e.g. the needles of *Shigella* sp. or *Yersinia* sp. are 50-60 nm in length, whereas the needle of pathogenic *E. coli* is 600 nm in length (31, 64). Signaling from the host membrane pore via the needle to the base structure after cell-cell contact triggers the secretion mechanism, which is energized by ATP hydrolysis and the PMF (65-67). The substrate is targeted via an N-terminal secretion signal together with a chaperon in a chaperon-effector-complex to the secretion system. The chaperone-effector-complex is then bound by an ATPase which unfolds the effector protein and triggers it to the cup-like structure at the inner membrane (68). The function of the chaperone is still not fully understood and it is assumed that it could be involved in several functions: i) keeping the effector in an unfolded inactive state, ii) triggering the secretion process, or iii) preventing an association of the effector with the membrane (33, 60). T3SSs are present in a broad range of pathogens like *Yersinia enterolytica*, *Salmonella enterica*, *Salmonella typhimurium*, *Shigella flexneri*, *Burkholderia pseudomallei* or *E. coli* (60).

Type V secretion systems (T5SS) are the simplest secretion systems within bacteria and are widely spread among the group of Gram-negative bacteria (33). Secretion in T5SSs occurs in a cell-cell contact dependent manner and enables the transfer of large proteins (31). The secreted proteins, which are often virulence factors, either remain bound to the cell surface (e.g. adhesins) or are cleaved by an extracellular protease (e.g. toxins) (69). To date two different types of T5SSs are described in detail; autotransporters and 2-partner secretion systems (69, 70). Autotransporters are mostly used for the secretion of toxins or adhesins, whereas 2-partner secretion systems are used to co-ordinate the formation of biofilms or to compete on environmental resources (e.g. via growth inhibition) (71-73). The secretion mechanism is a SecYEG dependent 2-step mechanism. Autotransporters are targeted with a signal peptide to the inner membrane and are secreted into the periplasm via the Sec-translocase. The C-terminal domains of autotransporters integrate into the outer membrane forming β-barrels to enable the secretion of their own N-terminal domains. Therefore, autotransporters consist of three different domains: a C-terminal domain for β-barrel assembly, an N-terminal domain with a signal peptide, and a passenger domain in between (69, 74, 75). The 2-partner secretion systems are very similar to autotransporter systems but consist of two proteins. One protein is forms the outer membrane β-barrel and the other protein is secreted across the membrane. The translocation of both proteins into the periplasm is SecYEG dependent. The effector protein is first secreted as a pre-pro-protein into the periplasm where it is cleaved into a pro-protein prior transfer through the outer membrane. Outside the cell the pro-protein is then cleaved to the active protein (76-78). Often 2-partner systems are involved in contact dependent growth inhibition (CDI) in which the secreted effector protein leads to cell-death of the recipient by inhibiting the PMF (e.g. *E. coli*, *B. pseudomallei*, or *Yersinia pestis*) (79, 80). Another example for T5S is the secretion of the IgA protease in *N. gonorrhoeae* and *N. meningitides* (81).

Type VI secretion systems (T6SS) are evolutionary related to the DNA injection systems of bacteriophages and share several proteins with high sequence identity and structural similarities with them (82). Many Gram-negative bacteria encode T6SSs involved in virulence or self-recognition within the same species (e.g. *V. cholera*, *P. aeruginosa*, *A. tumefaciens*, *Edwardsiella tarda*, *Burkholderia mallei*, or *Proteus mirabilis*) (83-86). T6SSs consist of multiple proteins and assemble to a basal plate located in the cytoplasmic membrane, an injection-tube that is pushed outside the cell and initiates the cell-contact, and a hexameric AAA^+ ATPase (87). The detailed mechanism of T6S is still not fully understood and needs to be elucidated further. Up to now it is not clear if T6SSs have additional proteins with lysozyme-like activity at the tip of the tube to achieve penetration of the

host membrane (31). At present the proposed mechanism of T6S seems to be in a way similar to that of T3SSs.

The most recently identified class of type VII secretion systems (T7SS) can be found in *Mycobacteria*, as well as in several other Gram-positive bacteria. Many pathogenic *Mycobacteria* use T7SSs to secret virulence factors (e.g. *Mycobacterium marinum*, *Mycobacterium tuberculosis*, or *Mycobacterium leprae*), but within *Mycobacterium smegmatis* T7SSs are also used for conjugation (88-91). Conserved gene clusters related to T7SSs (WXG100 family or ESX-4 homologues) were also identified in other Gram-positive bacteria like *Bacillus anthracis*, *Bacillus subtilis*, *Staphylococcus aureus*, *Streptococcus agalactiae*, *Streptomycetes*, *Corynebacteria*, and *Nocardia* (32, 92-97). The first T7SS was identified in a *Mycobacterium bovis* BCG vaccine strain, which is used in the production of tuberculosis vaccine. While generating a non-virulent strain for vaccination a conserved ESX-1 gene cluster has been identified and demonstrated to be important for virulence of *M. bovis* (32). Up to now five ESX clusters are identified and known to be involved in T7S within *Mycobacteria*. The ESX-1 and ESX-5 clusters are important for virulence, and in *M. smegmatis* ESX-1 is involved in conjugation. The functions of ESX-2 and ESX-4 are still not known and the ESX-3 cluster is involved in iron and zinc homeostasis (94). To date the mechanism of T7S is only poorly understood and a first model for the assembly of the secretion system and the secretion process is based on bioinformatical predictions and a few protein-protein interaction studies. T7SSs are multiprotein systems (ESX-1 encodes already 14 proteins) with an inner membrane channel associated with an inactive chaperone-like ATPase that interacts with the C-terminus of the effector proteins to form an active ATPase for secretion (32). It is assumed that the mycomembrane is crossed via a channel-like structure similar to other secretion systems, but so far no mycomembrane-channel components could be identified (32). Interestingly, the effector proteins of T7SSs dimerize prior secretion and are transferred as dimer-complexes (98, 99). Within these dimeric effector-protein-complexes one of the two proteins carries a C-terminal signal sequence (100). This signal sequence can be used for the secretion of other soluble proteins. Therefore, it is assumed that T7SSs secret multimeric complexes (101). Since no structural data is available and only a few proteins within T7SSs are characterized in detail, it is up to now not known how these complexes pass the inner membrane channel. However, T7SSs show some similarities to T4SSs (102, 103).

1.3 Multifunctional type IV secretion systems

The versatile and widespread group of type IV secretion systems (T4SS) can be separated into three subfamilies: i) the large group of conjugative systems, ii) the pathogenic relevant effector translocator systems, and iii) the small group of DNA uptake and release systems (104). Generally, T4SSs are large macromolecular machineries consisting of 11-15 different proteins (102, 104). These complex secretion machineries span both cell membranes in Gram-nagative bacteria and enable the secretion of single stranded transfer DNA and/or effector molecules to a host organism or into the environment, often in a cell contact dependent mechanism (104, 105). T4SSs are ubiquitous distributed among bacteria and archaea. Especially the large family of conjugative T4SSs is widely spread and can be found in Gram-negative, Gram-positive, and wall-less bacteria, as well as in Crenarchaeota (105). However, the subfamilies of effector translocator systems and DNA uptake/release systems can only be found in Gram-negative bacteria (104, 105). The rife of T4SSs

results in considerable genome plasticity, which has a strong impact on the rapid evolution of bacterial genomes (105). The presence of a T4SS is often coupled with the acquisition of different environmental or metabolic advantages to its host. For instance, T4SSs are major initiators for increasing antibiotic resistances or enhance the virulence and colonization skills of pathogenic bacteria (106-108).

The broad variety of T4SSs hampers a general classification of these systems. Up to now, three different classification methods are commonly used and are currently under discussion. The simplest and most limited method is the classification of T4SSs into the three previously mentioned subfamilies (104). Considering the diversity of conjugative plasmids grouped together within the family of conjugative systems, a detailed sub-grouping of the different conjugative plasmids was established with the classification method based on the incompatibility groups of the plasmids (30, 109). A new T4SS classification method based on sequence similarities and biochemical data obtained from relaxases was established by Garcillán-Barcia *et al*. (110), and represents in combination with the other two classification methods a logical and reasonable approach for the difficult classification and grouping of T4SSs. Recently, this classification system was extended to a separation system based on the differentiation of eight relaxase families and eight mating pair formation (MPF) families (111, 112).

The conjugative systems are the best studied subfamily of T4SSs and are distributed among self-transmissible plasmids, ICEs and GIs (30, 113-115). The *E. coli* F-plasmid and the *A. tumefaciens* Ti-plasmid T4SSs represent only two well characterized prototypes of this large subfamily (30, 102, 116). Functionally, conjugative T4SSs transfer ssDNA form one organism to another in a cell contact dependent mechanism (102). Thereby, the conjugative DNA transfer is not limited to one species and can occur within and between kingdoms (104, 107). The mechanism can be divided into three reactions types: i) the initial DNA processing reaction, ii) the recruitment of the substrate to the secretion complex, and iii) the translocation process (102, 117, 118). The initial DNA processing reaction is of major interest within this study and will be described in detail in Chapters 1.7 – 1.9. The substrate recruitment is often mediated by positively charged C-terminal signal-sequences and is covered in Chapters 1.7 and 1.8. The translocation process of T4SSs still includes many unknown variables and different translocation models are currently under discussion. Recently, Fronzes *et al.* solved the structure of an assembled and functional purified secretion complex, which represents a milestone in understanding the translocation process in T4SSs (119, 120). Detailed information about the assembly of this representative secretion complex will be given later within this Chapter in context with the functional mechanism of T4SSs.

The family of effector translocator systems is functionally closely related to the conjugative systems. In addition to the often coupled conjugative DNA transfer, these systems transfer effector proteins (e.g. virulence factors, toxins, or metabolically active substrates) into host cells in a cell contact dependent mechanism, similar to that of conjugative systems (102, 104, 121-123). The transferred effector proteins and the T-DNA play an important role in infection processes, colonization processes, and the survival of bacteria within the host cell or on tissues (102, 104). These systems often lack *dtr* proteins and the coupling protein, indicating their high specialization on effector proteins transfer (105, 122, 124). The substrate recruitment is often mediated by Chaperones or C-terminal recognition signals, similar to conjugative systems (105). The mechanism of the effector translocation through the secretion complex still remains unknown and is topic of current research.

The small subfamily of DNA uptake and release systems resembles up to now only three known examples, which are the DNA uptake systems of *Campylobacter jejuni* (Cjp/VirB) and *Helicobacter pylori* (ComB), and the DNA release system encoded in a genomic island of *N. gonorrhoeae* (GGI) which is in the major focus of this study (113, 125, 126). The remarkable and novel cell contact independent T4SS of *N. gonorrhoeae* will be described in detail in Chapter 1.6.

To enable the transfer of DNA and effector proteins from one organism to another, T4SSs have to pass the barrier of the inner and the outer cell membrane and the border of the host cell membrane. Therefore, T4SSs resemble a huge macromolecular complex spanning both membranes. The coupling protein (e.g. $VirD4_{AT}$, $TraD_{F}$, $TrwB_{R388}$), which mediates the contact between the substrate and the secretion complex is integrated into the inner membrane and localize in close proximity to the secretion complex (detailed information on T4SS coupling proteins is given in Chapter 1.10). Besides the secretion channel and the coupling protein, T4SSs encode one or two additional ATPases thought to energize the assembly of the secretion complex, the pilus assembly and disassembly, and/or the substrate transfer (102, 105). $VirB4_{AT}$ like ATPases (e.g. $TrwK_{R388}$, $TraC_{F}$, $TrhC_{R27}$) can be found in all T4SSs and have been reported to be involved in pilus assembly ($TraC_{F}$) (127). It has been proposed that these proteins assemble into an active homohexameric state and are involved in substrate docking (105, 128). Recently the structure of the ATPase domain of the VirB4 homologue of *Thermoanaerobacter pseudethanolicus* has been solved and has been modeled into the structure of the T4SS core complex (129). This protein forms an equilibrium of monomers and dimers in solution. Walldén et al. also demonstrated that the VirB4 homologue of plasmid pKM101 binds to the core complex as a monomer (129).

$VirB11_{AT}$ like ATPases (e.g. $TrbB_{RP4}$, $TrwC_{R388}$) have been reported to assemble as homohexameric rings and associate with the cytoplasmic membrane (105). The activity of these proteins seems to be enhanced *in vitro* in the presence of lipids (130, 131). Homologues of $VirB11_{AT}$ like ATPases are often found in type IV pilus systems, type II secretion systems and archaealla (105, 132). In T4SSs these proteins are essential for the assembly of the pilus and the secretion complex (105). Remarkably, no VirB11 homologues can be found in T4SSs that resemble the F-plasmid. The membrane spanning secretion complex consists of an inner membrane channel and an outer membrane channel that are connected in the periplasmic space. The inner channel is made of $VirB3_{AT}$, $VirB6_{AT}$, $VirB8_{AT}$, and $VirB10_{AT}$ like proteins, which tightly interacts with each other to form a stable channel complex (102, 105, 119). $VirB3_{AT}$ like proteins are integral inner membrane proteins and have been reported to be involved in pilus assembly and substrate translocation, even though the specific function of these proteins is still unknown (105, 133, 134). $VirB6_{AT}$ like proteins (e.g. $TraG_{F}$) are polytopic membrane proteins that stabilize the *mpf* complex and mediate the transfer of the ssDNA via direct DNA-protein interactions (135-138). $VirB8_{AT}$ like proteins are responsible for the spatial positioning of the secretion complex at the cell pole and have been reported to form dimers (105, 139, 140). $VirB10_{AT}$ homologues are bitopic inner membrane proteins with a structurally dynamic character (141). These proteins are structural scaffold proteins and bridge the inner and outer channel components in a structural dynamic mechanism, including conformational changes due to energizing via ATP-hydrolysis (119, 120, 142). The periplasmic connection and the outer membrane channel consist of $VirB1_{AT}$, $VirB2_{AT}$, $VirB5_{AT}$, $VirB7_{AT}$, and $VirB9_{AT}$ like proteins (102, 105, 119). $VirB1_{AT}$ like proteins are transglycosylases and do not directly interact and associate with the secretion complex, but are essential to facilitate the assembly of the secretion channel via local degradation of peptidoglycan (143). The outer membrane channel is mostly assembled of $VirB7_{AT}$ like proteins (e.g. $TraO_{pKM101}$),

which are small lipoproteins that stabilize $VirB9_{AT}$ like proteins (105, 144). $VirB7_{AT}$ like proteins form together with $VirB9_{AT}$ and $VirB10_{AT}$ like proteins the central membrane spanning channel (105, 119, 120). The $VirB9_{AT}$ like proteins (e.g. $TraN_{pKM101}$) are part of the outer membrane channel and are hydrophilic proteins that localize in the periplasm and the outer membrane (105, 145). The cell contact and substrate transfer in T4SSs is facilitated by the formation of a dynamic pilus structure at the outside of the cell, which is formed by $VirB2_{AT}$ and $VirB5_{AT}$ like pilin subunits. $VirB2_{AT}$ like proteins are small hydrophobic channel components that could not be readily identified in Gram-positive bacteria and archaea (105). $VirB5_{AT}$ like proteins are transferred into the periplasmic space and facilitate the transfer of the substrate and the assembly of the pilus (105). $VirB5_{AT}$ like proteins have been reported to mediate the cell contact with the host (105).

The number of available structures of T4SS components is strongly increasing within the last years, giving more and more insight into the complex mechanism of substrate transfer and complex assembly. Up to now, structures of $VirB5_{AT}$ like pilin subunits, $VirB11_{AT}$ like ATPases, $VirB7_{AT}$ – $VirB10_{AT}$ like channel components, and $VirD4_{AT}$ like coupling proteins are known (119, 120, 146). These structures allow in combination with biochemical data limited conclusions about a common architecture model of T4SSs. Recently, the first multi-protein structure of the 1 MDa T4SS core complex from plasmid pKM101 was solved with a resolution of 15 Å, using cryo electron microscopy and 3D crystallography (119, 120). This remarkable breakthrough allowed for the first time structural insights into the assembly and functionality of a membrane spanning multi-protein complex. The secretion complex of pKM101 has an overall height of 185 Å and consists of an inner layer (I-layer), an outer layer (O-layer) and a cap like structure (119). The I-layer forms a double-walled ring-like chamber with a 55 Å central opening at the cytoplasmic site and is made of the N-terminal domains of $TraO_{pKM101}$ ($VirB9_{AT}$) and $TraF_{pKM101}$ ($VirB10_{AT}$) (119). The O-layer resembles together with the cap the main body of the secretion complex and is formed by $TraN_{pKM101}$ ($VirB7_{AT}$) and the C-terminal domains of $TraO_{pKM101}$ ($VirB9_{AT}$) and $TraF_{pKM101}$ ($VirB10_{AT}$) (119). The cap has a central constriction of 10 Å in diameter, representing the outer membrane opening of the secretion channel (119, 120). The constriction can be extended to a diameter of 20 Å via conformational rearrangements and allows a controlled opening and closing of the secretion channel (119, 120).

However, the exact mechanism of substrate transfer through the secretion channel (especially the transfer of large proteins) still remains largely unknown. Two different models for the substrate transfer mechanism in T4SSs have been proposed. In the channel model the pilus is supposed to form a channel connecting the host cell with the recipient cell (104). The extension of the pilus mediates the initial cell contact and the followed retraction of the pilus brings both cells into close proximity to each other, allowing an efficient substrate transfer through the secretion channel (104). In the piston model the substrate is loaded on top of the pilus and is transferred to the recipient cell during the extension of the pilus. The secretion channel acts thereby like a piston motor and pushes the substrate to the outside (104). Nevertheless, both models still raise the question of how the substrate enters the secretion channel. Especially the mechanism of the exceptional secretion mechanism of the pertussis toxin from *Bordetella pertussis*, which is mediated by the *sec* pathway and the T4SS, represents an interesting model to study the substrate entrance into the secretion complexes of T4SSs.

1.4 Secretion Systems of *Neisseria* sp.

Genome analyses of different *N. meningitides* strains, *N. gonorrhoeae* strain FA1090, and a strain of the commensal *Neisseria lactamica* revealed that *Neisseriacea* contain at least three different types of secretion systems (T1SS, T4SS and T5SS) and uncovered a strongly degenerated secretome in *N. gonorrhoeae* in comparison with *N. meningitides* and *N. lactamica*. Many of the pseudogenes and disrupted ORFs in *N. gonorrhoeae* encode genes of different secretion systems (147). The degenerated secretom of *N. gonorrhoeae* should be considered with caution since only the genome sequence of strain FA1090 was available by the time of the analysis (147). Especially due to the adaptation to the specific niche *N. gonorrhoeae* is inhabiting it is reasonable, that the number of secretion systems within this species is reduced. In this comparative genome analysis no homologous genes for T2SS, T3SS, T6SS, and T7SS were identified (147). However, it cannot be excluded that still unknown or alternative secretion mechanisms like e.g. the transport via outer membrane vesicles in *E. coli* (148) are present in *N. gonorrhoeae*. The latter option is particularly noteworthy since *Neisseria* are known for an intense blebbing activity (14). Homologues for T1SS genes can be found in *N. meningitides*, but not in *N. gonorrhoeae* and *N. lactamica*. *N. meningitides* encodes three RXT homologous islands (RTX I-III) encoding *frp* (Fe-regulated proteins) genes, which are expressed under iron starvation (149). The exact function of these proteins still needs to be investigated. *N. meningitides* encodes besides the *frp* gene cluster two other T1SSs (*natC* and *mtrE*), which show similarity with multidrug efflux systems (147). Some strains of pathogenic *Neisseria* contain contrary to commensal *Neisseria* a T4SS, which is remarkable in its ability to secret DNA into the environment. At first this T4SS was only identified in *N. gonorrhoeae* strains, but recently it was also identified in a few *N. meningitides* strains (113, 150). A detailed description of the neisserial T4SS is given in Chapter 1.6. The most abundant secretion systems in *Neisseriaceae* are T5SSs with a broad range of versatility (e.g. *autA/B*, *ata1/2/3*, *iga*, *app*, *nhhA*, *ihhA/B*, *nalP*, *nadA*, *aus1*). At least three of these autotransporter gene clusters (*iga*, *app,* and *nalP*) seem to be highly conserved among *Neisseriaceae* (147). The IgA protease is absent in commensal *Neisseria* but is very important for the virulence of pathogenic *Neisseria*, as it cleaves human IgA, Lamp1 and chorionic gonadotropin (151-153). The adhesion and penetration protein App is conserved in *N. meningitides*, *N. gonorrhoeae*, and *N. lactamica* and is involved in the adhesion to distinct epithelial surfaces (154, 155). The NalP protein contains an uncommon C-terminal lipoprotein motif. The lipidation of NalP is important for the positioning of the protease near the cell surface to enable a proper access to the proteolytic sites of secreted proteins (156).

1.5 Mobile genetic elements

The spread of genetic information via horizontal gene transfer is one of most important forces of bacterial evolution. Three different mechanisms are known to be used for DNA transfer in bacteria: i) transduction via phage infection, ii) natural transformation via competence systems, and iii) T4SS mediated conjugation (111, 157). Since conjugation can result in the transfer of whole genomes it has the highest impact on DNA transfer in bacteria (158). T4SSs are not only located in conjugative plasmids, besides they can be also found within bacterial genomes located in or next to self-transmissible conjugative elements, or as T4SSs that are no longer involved in DNA transfer but in secretion of effector proteins (159).

Within bacterial genomes several genetic elements are self-transmissible and are classified as ICEs (integrative conjugative elements). This includes also conjugative transposons and integrated conjugative plasmids (e.g. Tn*916* or CTnDOT with low sequence specificity or more sequence restricted ICEs like ICE-SXT) (160-162), or GIs (genetic islands) containing a compatible *oriT* sequence, or GIs that are integrated into a conjugative element (e.g. ICEclc or ICEM1Sym) (159). It is assumed that non mobile GIs are evolutionary related to ICEs but underwent a genetic degradation resulting in a loss of mobility (159). Besides ICEs, also integrative mobilizable elements (IMEs) are present in bacterial genomes. Theses genetic elements need the presence of other integrated T4SSs to be mobilized, and differ in this point from ICEs, conjugative transposons and GIs (163, 164).

ICEs are wide spread within bacterial genomes and an extensive bioinformatical analysis to classify and identify different groups of ICEs by the use of chromosomally integrated T4SSs was done in 2011 by Guglielmini *et al.*(111). The classification followed from this analysis is based on the comparison of the different relaxases families (MOB_B, MOB_T, MOB_V, MOB_Q, MOB_P, MOB_H, MOB_F, MOB_C), coupling proteins, conserved VirB4 like ATPases, and eight different MPF groups (MPF_T, MPF_F, MPF_G, MPF_I, MPF_B, MPF_C, MPF_{FA} and MPF_{FATA}) (111, 112). The analysis revealed that all glades of bacteria with sequenced genomes contain at least 86% more ICEs than conjugative plasmids. The distribution of ICEs and conjugative plasmids can vary strongly within different glades and both systems can co-occur sometimes in the same organism in large numbers. This implies that different T4SSs can co-exist in a functional state within one host. Several ICEs combining different MOB and MPF systems were identified in the analysis of Guglielmini *et al*., indicating an evolutionary shuttle within these systems (e.g. MPF_T + MOB_P, MPF_G + MOB_F or MPF_G + MOB_H) (111).

Three different modules enable DNA transfer in ICEs: i) an integration and excision module, ii) a conjugation module, and iii) a regulatory module. In addition to these modules ICEs contain a broad range of extra functions like antibiotic resistances, heavy metal resistances, nitrogen fixation, degradation of aromatic compounds, virulence, and biofilm formation. Due to the spread of antibiotic resistance and the high number of mobile pathogenicity islands (e.g. 89K from *Streptococcus suis*) ICEs are of important clinical relevance (165-171).

Before ICEs can be transferred via T4SSs to a recipient cell, the integrated element needs to be excised from the chromosome to form a circular extra-chromosomal intermediate. These intermediates are normally not auto-replicative and cannot be maintained in an extra-chromosomal stage. The excision within ICEs is mediated by integrases (often phage-like integrases), which are proteins mostly belonging to the family of tyrosine recombinases (e.g. Int from phage λ (172), SXT-R391 (162), ICEclc or ICEM1Sym (173, 174)), but also serine recombinases (e.g. Tn5397 (175)), DDE transposases (e.g. CTnGBS2 (176)), or XerCD like integrases (e.g. PAPI-1 or ICEHin1056 (177, 178)) can be found. The mechanism of excision is comparable to the mechanism of integration (site-directed recombination) and the integrase is forced to excise the ICEs via additional recombination directionality factors (RDFs). RDFs (excisionases) are small DNA-binding proteins that interfere with the DNA or perform protein-protein interaction and are very difficult to identify with bioinformatical approaches (179). Up to now, only a few RDFs are known and have been functionally characterized (e.g. Exc from CTnDOT or Xis from 89K (180, 181)). It is demonstrated that RDFs can show Topoisomerase I like DNA processing activity *in vitro* (180). Both proteins, the integrase and the RDF, are mostly not co-regulated since the frequency of excision and integration needs to be controlled separately due to environmental influences. An excellent review on the regulation of ICEs is given by

Wozniak *et al.* (159). The excision of ICEs can lead to the loss of the genetic elements. Up to now two mechanisms are known to prevent the loss of the excised ICEs. One, still uncharacterized mechanism is the stabilization of the excised ICEs with the help of additional proteins involved in chromosome partitioning and plasmid maintenance (e.g. Soj from PAPI-1 (182)), the second mechanism is the stabilization of excised ICEs via toxin-antitoxin systems (e.g. ICE-SXT (183)).

For conjugational DNA transfer ICEs need at least a minimal set of T4SS components, which are a relaxase for the processing of single stranded transfer DNA, a coupling protein to mediate the contact between the DNA and the secretion channel, and components of the mating pair formation complex (MPF) (111). For many ICEs a relaxase and an *oriT* sequence are identified and are located in close proximity to each other. One exception is given by the relaxase TraI of ICE-SXT that has a distance of 40 kb to its *oriT* sequence (184). Recently, the bioinformatical analysis of Guglielmini *et al.* revealed several T4SSs within ICEs lacking the relaxase protein. The authors explained this result with a high number of potentially unknown relaxases that are not recognized within the analysis and a high number of genomic T4SSs that are not involved in conjugation, but might be involved in transfer of other substrates (111). Besides the use of T4SSs for DNA transfer a few ICEs mediate the DNA transfer via DNA translocases similar to the mechanism of SpoIIIE or FtsK (185).

Likewise conjugative plasmids, ICEs also show in part exclusion groups to prevent the acquisition of closely related elements in one host (186, 187). To prevent a double-integration of the same ICE in the chromosome bacteria use a similar mechanism for entry exclusion like it is used for conjugative plasmids. The entry exclusion is mediated via the coupling protein of the T4SS and an additional entry exclusion protein. After the successful transfer of an ICE to the recipient cell it is integrated into the recipient chromosomes on integrase-dependent integration sites (e.g. *dif*-sites or *attR*- and *attL*-sites (159, 188, 189)). Many integrases use a mechanism similar to Topoisomerase I activity and do not require additional energy from ATP hydrolysis (172). The site-specific recombination mechanism of integrases is comparable with the mechanism of DNA integration of phages (e.g. phage ϕ80 and phage P4 (190, 191)). If the primary insertion site is not present in the chromosome ICEs can often also integrate on a second, partly imperfect insertion site (192).

1.6 Gonococcal Genetic Island

The Gonococcal genetic island (GGI) is a horizontally acquired genetic island with a size of 57 kb and takes its name from gonococci since it was first identified in the chromosome of *N. gonorrhoeae* (113). It is present in 80% of *N. gonorrhoeae* strains and in 17.5% of *N. meningitides* strains (only in serogroups H, W-135, Z and strain α275), and so far it could not be detected in commensal *Neisseria* species (e.g. *N. lactamica*) (113, 150, 193, 194). Within different neisserial strains the genetic organization and the integrity of the GGI can vary and five different GGI classes can be distinguished (150, 194). An overview of the different GGI classes is given in Table 1. Some variants of the GGI can be linked with the incidence of disseminated gonococcal infections (DGIs), but the presence of the GGI is in general not necessarily required for pathogenicity within *Neisseria* (113). The GGI variants isolated from DGI patients mostly contain an additional serum-resistence allele (*sac-4*) located within the *traG* gene and the lytic transglycosylase *atlA* (113, 195). However, it is not known if the product of the expressed *sac-4* allele is surface exposed, resulting in a serum-resistance (113). In

N. meningitides no correlation between the presence of the GGI and the invasive form of the disease could be shown (150). Zola *et al.* showed recently a coherency of the GGI encoded T4SS with the ability for intracellular survival of *N. gonorrhoeae* in absence of the TonB dependent iron acquisition complex (196).

Table 1: Overview of different GGI classes found in *N. gonorrhoeae* and *N. meningitides* strains. The data is summarized from Woodhams *et al*. (150).

Class I	lacks *atlA* and *exp1* deletion of 11 kb between *yecB* and *parB* encodes *eppA* long *traA* allele 46 kb in size
Class II	lacks *atlA* and *exp1* deletion of 8.2 kb between *yeb* and *ssb* but integrated IS1301 element encodes *eppA* insertion of IS1301 element in *traD* with second copy in *yaa* long *traA* allele 51 kb in size
Class III	lacks *atlA* and *exp1* deletion of 300 bps in *traD* and deletion of *yaa* insertion of IS1665 element in *traK* (no TraK production possible) 56 kb in size
Class IV	highly similar to Class III within *difA* and *ydbA* numerous insertions within *ybdB* and *parA* deletions of 56 bps in *yeh*, 80 bps in *topB*, 30 bps in *parB* unique for *N. meningitides* strain α275 64 kb in size
Class V	deletion of 22 kb between *traE* and *yda* (replaced with IS1301 insertions) deletion of 11 kb between *yecB* and *parB* contain *ssbB* and a homolog of *topB* 24 kb in size

Based on several differences between the neisserial chromosome and the GGI it was proposed that the GGI was horizontally acquired. These differences are a variation within the GC content (44% in GGI vs. 51% in chromosome in *N. gonorrhoeae*), as well as differing dinucleotide frequencies between the GGI and the chromosome, and diverging quantities of DUS sequences which are necessary for successful transformation of *Neisseria* species (in the chromosome 1.1 DUS/1kb vs. 9.5 DUS/1kb in the GGI of *N. gonorrhoeae*) (197-199). Besides the already referred evidences for a putative horizontal acquisition of the GGI, it is bordered by two *dif* sites (*difA* and *difB*) which are recognized by the genome integrated XerCD recombinase (189, 198). The integration of the GGI into the chromosome could be facilitated via a XerCD dependent site-directed recombination mechanism (189, 200). The excision frequency of the GGI is very low due to a stabilizing effect of the imperfect *difB* site (189). *DifB* contains four mismatches compared to the consensus sequence of the intact *difA*

site and reduce therefore the excision efficiency of XerCD (189, 200). The circular form of the GGI cannot be maintained sufficiently in the cell, which could lead to the loss of the GGI since it seems not to be self replicative (189, 198, 200). All sequenced gonococcal GGIs contain the same four mismatches in the *difB* site, whereas in meningococci the mismatches in *difB* differ, which explains the chromosomal maintenance of incomplete GGI versions in meningococci (189). All these findings lead to the conclusion that the GGI was horizontally acquired from an up to now unknown ancestral species (200, 201). Most likely the GGI is an ICE that underwent genetic degradation over time, resulting in becoming a stabilized genomic island (198, 200). This assumption is consistent with the finding that the GGI encodes a T4SS with a large set of IncF and IncH homologous genes (110, 113, 198, 202).

The GGI of *N. gonorrhoeae* strain MS11 contains in total 62 open reading frames (ORFs) organized in at least five different operons ((113, 198, 203), K. Siewering unpublished data). The first 27.5 kb of the GGI encode a unique T4SS that enables gonococci to secrete ssDNA directly into the environment in a cell-contact independent manner (198). The secreted DNA is 5' protected against DNAses and can be used for natural transformation by other gonococci and, therefore, is a potent factor for the spread of genetic information (202). It is unclear whether or not the secreted DNA also has an immunoregulatory function within the immune defense of gonococci (113, 204). Within meningococci the T4SS neither is involved in DNA secretion nor in a Ton-independent iron uptake for intracellular survival like it has been reported for *N. gonorrhoeae* (150, 196).

The T4SS is encoded within the first three operons of the GGI and consists of 23 proteins ((198), K. Siewering unpublished data). 18 genes are homologous to the *E. coli* F-plasmid (IncF) (e.g. $traG_{NG}$ and $traG_{EC}$ have 22% identity and 40% similarity; $traH_{NG}$ and $traH_{EC}$ have 24% identity and 42% similarity; $traF_{NG}$ and $traF_{EC}$ have 29% identity and 45% similarity) (113, 205, 206). Most of the proteins are involved in the formation and stabilization of the secretion complex. 15 of the genes are ordered in the same orientation like the *E. coli* F-plasmid (198). The genes for the assembly and the stabilization of the secretion complex are located within the second and the third operon ((198), K. Siewering unpublished data). A few genes of the T4SS differ from those of the F-plasmid, e.g. the peptidoglycan transgylcosylase LtgX, the relaxase TraI, the coupling protein TraD, as well as the small proteins Yaf and Yaa (198, 202). However, the DNA processing and transferring proteins (TraI and TraD) of the neisserial T4SS seem to be phylogenetically more related to those of the IncH group and are located together in the first operon with the small proteins Yaa and Yaf ((110, 198, 202), PhD thesis of E. Pachulec 2010). The relaxase TraI belongs to the family of MOB_H relaxases typically associated with ICEs and GIs, as well as conjugative plasmids (110). The coupling protein TraD clusters in phylogenetic analyses together with this group of relaxases and homologous proteins to Yaa. A predicted origin of transfer (*oriT*) is located within an inverted repeat in an AT-rich region between the coding sequences of *yaf* and *ltgX* in close proximity to the relaxase *traI* (150, 202). More detailed information considering the DNA processing proteins of the T4SS of *N. gonorrhoeae* are given in the Chapters 1.7 – 1.9. In addition to the genes of the T4SS the GGI encodes 39 ORFs with conserved hypothetical genes with up to now unknown function and some with putative DNA binding and processing function (e.g. the chromosome partitioning proteins ParA and ParB, the putative helicase Yea, Ssb, TopB and the DNA methylases Ydg and YdhA) (198). For most of the GGI encoded T4SS genes, DNA secretion assays were performed to test whether they are essential for DNA secretion or not ((113, 198, 202, 205), PhD thesis of E. Pachulec 2010). A scheme of the genetic organization of the GGI and an overview about the essential genes for DNA secretion is given in

Figure 3. Nearly all genes involved in assembly and stabilization of the secretion complex have been shown to be essential for DNA secretion. In particular, mutations within TraD do not result in a complete prevention of DNA secretion, but led to a significantly decreased level of secretion (198). Currently the small protein Yaa within the neisserial GGI is not tested in secretion assays, but a homologous protein within SXT/R391 S043 has been shown to be essential for DNA transfer (207).

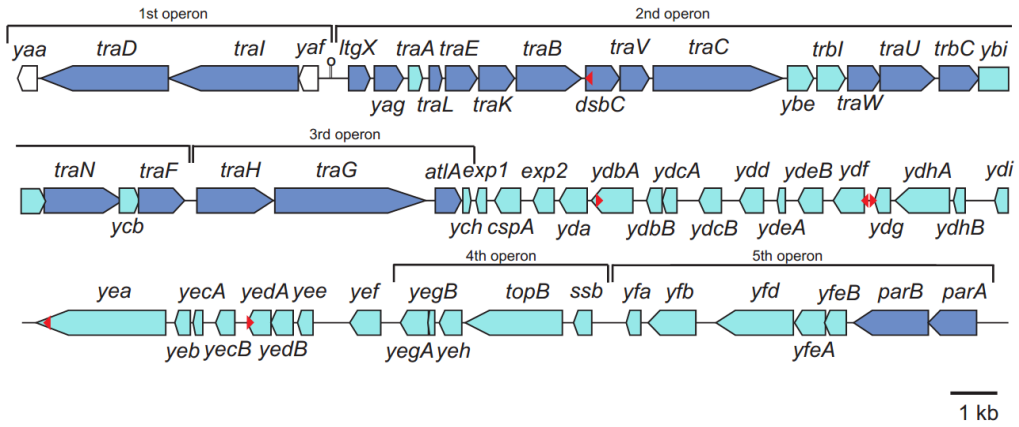


Figure 3: Genetic organization of the GGI from *N. gonorrhoeae* strain MS11. Genes shown to be essential for DNA secretion are colored in blue, genes shown to be not essential for DNA secretion are colored in cyan. Genes that have not yet been proven to be essential for DNA secretion are white colored. The *oriT* is located between *yaf* and *itgX*. The six DUS sequences present in the GGI are labeled with red triangles. The picture is adopted and changed according to the PhD thesis of E. Pachulec 2010, the master thesis of K. Siewering 2010, and (200, 203).

At present, there is not much known about the regulation of DNA secretion in the neisserial T4SS or the excision and integration of the GGI. So far, no homologues of regulatory factors could be identified in the GGI. Recently, Salgado-Pabòn *et al.* showed that the DNA secretion is increased in pilliated *N. gonorrhoeae* cells, resulting from an upregulation of the expression levels of *traI* and *traD* (208). This is the first indication of a regulatory mechanism within the T4SS of *N. gonorrhoeae*. Most likely the system is also regulated on other levels which have not yet been identified, e.g. regulatory mechanisms for the excision and integration of the GGI might exist. As discussed previously in Chapter 1.5, the excision of ICEs and GIs is often triggered by small proteins functioning as RDFs. Since RDFs are difficult to identify, it could be possible that within the hypothetical genes of the GGI putative RDFs are encoded. The control on excision and integration of the GGI would enable the cell to distinguish between the secretion of the excised GGI (similar to conjugative plasmids) and the secretion of the whole chromosome (like Hfr-strains in *E. coli*) under certain conditions.

1.7 DNA processing within type IV secretion

The DNA processing mechanism is one of the most important steps within bacterial conjugation and DNA secretion. The initial processing of double stranded supercoiled DNA into single stranded transfer DNA is a cascade like, highly structured mechanism and is essential to enable DNA transfer (209, 210).

Mostly, the DNA processing in T4SSs is mediated by a nucleo-protein-complex termed “relaxosome”, which consists of several proteins involved in DNA processing and the covalently bound transfer-DNA (209, 211, 212). The most important component of the relaxosome is the relaxase protein, which is responsible for the DNA cleavage and the processing of the transfer DNA (212-214). The sequence specificity and an optimal cleavage activity of the relaxase are mediated by additional nicking accessory proteins like $NikA_{R64}$, $TraJ_{RP4}$, $TraH_{RP4}$, $TraY_F$, $TraM_F$, or $TrwA_{R388}$, as well as the integration host factor (IHF) (209, 215-219). Some of them function as spatial positioning factors and trigger the relaxosome to the secretion complex (220), other NAPs introduce topological changes into the supercoiled DNA and enable access for the large relaxase protein (221).

The DNA processing mechanism is initialized by the binding of the NAPs and the IHF to distinct binding sites near the origin of transfer region (*oriT*) (214, 219, 222-224). The *oriT* sequence forms a stem-loop like structure resulting from an inverted repeat in an AT-rich region, which is recognized by the relaxase and other relaxosome components (225, 226). The inverted repeat is surrounded by various binding sites for NAPs, the IHF, and the relaxase, and could comprise sequence regions of several hundred bps (226). Generally, the *oriT* region of a T4SS is localized in close proximity to the *dtr* genes. The different NAP binding sites, as well as the binding site of the relaxase (*nic* site) in the *oriT* region are important triggering factors for the high sequence specificity of conjugative DNA processing (209, 219, 223, 227). The binding of dimeric NAPs to their distinct recognitions site enables a tetramerization of two NAP dimers, which leads to a bending of the supercoiled dsDNA (215). Detailed information about the functional mechanism of NAPs is given in Chapter 1.9. The heterodimeric IHF acts in a similar mechanism like other NAPs and bends the dsDNA helix around 160°, resulting together with other NAPs in a DNA bend that enables the relaxase binding-access to the dsDNA at the *nic* site (210, 228). Additionally to their DNA bending activity, NAPs interact directly with their cognate relaxase via protein-protein interactions and trigger the relaxase to its specific binding site (214, 218). Well studied systems for the relaxosome formation are the T4SSs of plasmid RP4 and of the F-plasmid (209, 211, 212, 214). For instance, the NAP $TraJ_{RP4}$ binds to its palindromic recognition site *srJ* forming a nucleo-protein complex, which interacts via protein-protein interactions with $TraI_{RP4}$, triggering the relaxase to its binding site *srI* near *nic* (214). It has been reported that the accessory protein $TraH_{RP4}$ stabilizes the $TraJ_{RP4}$-$TraI_{RP4}$-DNA complex, even though it does not directly bind to DNA (209). This sequential assembly of a relaxosome exemplifies the highly regulated initial process within conjugation and DNA secretion.

The mechanism of DNA processing and transfer in T4SSs is still under discussion and needs to be elucidated further. It is hypothesized that the relaxase protein binds in an inactive mode to the dsDNA and changes after a stimulus from other relaxosome and T4SSs components to its active mode (229). The assembled relaxosome is triggered to the coupling protein (CP) by spatial positioning factors (e.g. $TraM_F$) and is anchored to it via protein-protein interactions (220, 223). The CP oligomerizes to its active hexameric state after receiving an initiating signal, which comes most likely from the cell-cell contact with the recipient (229). This “pre-initiation” complex is connected to the secretion complex via the CP (229). A signal transduction cascade indicates the completed assembly of the secretion system and leads to a change in the binding mode of the relaxase, resulting in an activation of the DNA processing reaction (229). The nucleophilic attack of the active site tyrosine of the relaxase leads to a covalent phosphodiester-bond between the relaxase and the phosphate-backbone of the transfer DNA (212). More detailed information about the reaction mechanism of relaxases is given in Chapter 1.8. After the initial cleavage reaction of the relaxase

domain, the helicase domain of the relaxase starts to unwind the dsDNA in an ATP dependent mechanism. The single stranded transfer DNA stays covalently bound to the relaxase at its 5' end (212, 227). The CP starts to hydrolyze ATP and functions as a DNA pump, forcing the transfer of the DNA through the secretion complex (229, 230). The remaining uncleaved DNA strand is replicated by DNA polymerases via the rolling cycle replication mechanism (214). The DNA processing reaction is terminated by a second cleavage reaction at the *nic* site in the *oriT* region (212, 214, 231).

The DNA secretion system of *N. gonorrhoeae* represents an interesting exception of the common T4SSs since it secrets the DNA directly into the environment. The DNA processing mechanism of this T4SS is still unknown and the relaxase $TraI_{NG}$ belongs to a novel and uncharacterized relaxase family (MOB_H) (110). Besides the relaxase, no additional relaxosome components could be identified in *Neisseria*, even though a gene for the IHF is encoded within the chromosome. A possible role of the IHF in the formation of a putative relaxosome in *Neisseria* has not yet been investigated. A putative *oriT* region, which is located between the genes *yaf* and *ltgX* has been predicted in *N. gonorrhoeae* and mutagenesis studies in combination with complementation assays have shown that this region is essential for DNA secretion (202). Furthermore, the localization of the neisserial *oriT* is similar to that of *E. coli* F-plasmid, which is located between the relaxase gene *traI* and *geneX* (similar to the neisseria *ltgX*) (202). A possible role of the two small proteins Yaf and Yaa in the DNA processing mechanism in the neisserial T4SSs still need to be investigated. Considering the close proximity of both genes to the other *dtr* genes and their location within the same operon, both are important potential candidates for possible relaxosome components in *Neisseria*.

1.8 Relaxases

The most important step in conjugative DNA transfer is the processing of a single stranded transfer DNA. This crucial process is catalyzed by a single protein belonging to the diverse class of relaxases, which introduces a nick on the dsDNA in the *oriT* region (*nic* site) and unwinds the cleaved dsDNA to produce a single stranded transfer DNA (105, 212, 232, 233). This essential protein is part of the *dtr* machinery of T4SSs and its gene is located in close proximity to other *dtr* genes encoding additional relaxosome components and the CP (105).

The ubiquitous distribution and essential role of relaxases in conjugative DNA processing make these proteins an interesting target to develop a classification system for T4SSs. Based on sequence identities and biochemical data, Garcillán-Barcia *et al.* performed an extensive bioinformatical analysis and grouped the diverse class of relaxases into eight families (MOB_Q, MOB_P, MOB_V, MOB_F, MOB_H, MOB_C, MOB_{HEN} and Tn_{916} like relaxases) (110, 234). Considering the catalytic tyrosine(s) within the active site of the relaxases, the eight families can be divided into three categories. Relaxases with one catalytic tyrosine belong to the MOB_Q, MOB_P, and MOB_V families and can be clustered together into one major group (110). Relaxases with two catalytic tyrosines belong to the MOB_F family and differ in their enzymatic mechanism from the other relaxase families (110). The third group of relaxase families is made of MOB_H and MOB_C relaxases. For both families no biochemical data is available and their catalytic mechanisms are still unknown (110). An overview on the general characteristics of the six largest relaxase families is summarized from Garcillán-Barcia *et al.* in Table 2.

Relaxases fulfill several functions simultaneously to enable the conjugative DNA processing and transfer in T4SSs (212). Therefore, these proteins usually consist of at least two different domains, which are an N-terminal relaxase domain responsible for the DNA cleavage, a helicase domain that unwinds in an ATP dependent reaction the cleaved DNA strand from the 5' to the 3' direction, and sometimes an additional C-terminal domain (e.g. DUF1528) for protein-protein interactions that contains also a signal sequence to trigger the nucleo-protein complex to the secretion pore (103, 105, 212, 235). The composition of the distinct domains varies within different relaxases and strongly influences the mechanism of DNA processing. Biochemical data has revealed that the enzymatic activity of the helicase domain provides the required force that is necessary for a sufficient strand transfer (e.g. $TraI_F$) (212). However, not all relaxases contain a helicase domain (e.g. $TraI_{RP4}$) and use external replicative helicases to unwind the DNA (212). The relaxase domain is the only domain which is present in all relaxases and catalyzes the cleavage of the dsDNA at the *nic* site in the *oriT* region. The nicking reaction of relaxases is a site- and strand-specific transesterification reaction (212). The free energy released from the broken phosphoester-bond is restored in the covalent attachment of the 5' end of the DNA to the relaxase via a phosphotyrosyl-linkage. The resulting nucleo-protein complex forms a stable long-term intermediate (209, 236-238). The 3' end of the nicked transfer DNA strand remains bound to the relaxase during the unwinding process, keeping the superhelical topology of the DNA, but does not form a covalent bond with the protein (209, 236, 238). This bond is released after terminating the unwinding of the DNA and the single stranded T-DNA is further transferred via the secretion complex. The nicking reaction is terminated by a second cleavage event at the *nic* site (212, 214). Depending on the presence of one or two catalytic tyrosines, different possibilities for the termination of the reaction are in discussion. Relaxases with one catalytic tyrosine need a second relaxase for the termination of the reaction since the active tyrosine is blocked by the covalent phosphotyrosyl-linkage (212, 214). Structural data from $TraI_F$ indicate a dimerization of the protein supporting the hypothesis of a second relaxase for the termination, even though it is not clear whether this dimer formation is an artificial event during crystallization or the active oligomeric state of the protein (232). Biochemical and structural data tend to suppose a crystallization effect and not a natural dimer formation of $TraI_F$ (232). Nevertheless, dimerization has been reported for several other relaxases and strengthens the model for a possible termination reaction by a second relaxase (239, 240). Besides the dimerization model for relaxases with one catalytic tyrosine, the possibility of altered termination mechanisms need to be considered, since relaxases without a covalent intermediate have been reported (e.g. $MobC_{ColDF13}$) (241). These relaxases could terminate the DNA processing reaction by a second cleavage event performed at the same active tyrosine catalyzing already the first nicking reaction, implicating that these relaxases do not obligatory need a second protein. Generally, relaxases with two conserved tyrosines do not necessarily need a second protein for the termination reaction, although it is still unknown for many of them whether both tyrosines have a catalytic activity (212).

The reversible and isoenergetic nicking reaction of relaxases is enabled by three distinct sequence motifs, which are highly conserved in almost all relaxase families (110, 212). However, some relaxase families have altered sequence motifs. The influence of these changes on the DNA processing reaction still remains unknown. The first motif (Motif I) consists of one or two conserved tyrosines, with one of them acting with its hydroxyl-group as a nucleophil on the phosphate-backbone of the DNA (242). The catalytic tyrosine attacks the phosphate-backbone at a distinct scissile phosphate at the *nic* site and forms a covalent phosphodiester-bond with the cleaved transfer DNA strand (Y-O-P-O-X) (212). The remaining strand contains an OH-group at its 3' end. The second motif

(Motif II) contains a conserved aspartate or glutamate and is thought to maintain the protein-DNA contact at the 3' site of the nick at *oriT* and enhances the nucleophilic attack of the catalytic tyrosine (110, 242, 243). The third sequence motif (Motif III) contains several conserved histidines involved in the coordination of a divalent metal ion, which is in the main Mg^{2+} (110, 242, 243). The coordinated metal ion is in close proximity to the catalytic tyrosine and enables the nucleophilic attack of the hydroxyl-group of the tyrosine by pulling electrons from the oxygen and forcing the OH-group to act as a nucleophil (212, 232, 244). This highly conserved motif is altered within the MOB_H family and seems to be a special characteristic for these relaxases (110). The altered histidine triad and conserved HD domains in MOB_H relaxases suggest a different metal ion coordination mechanism of this relaxase family (110). How the metal ion is coordinated in MOB_H relaxases and how this changed mechanism influences the cleavage mechanism remains unknown, since no biochemical data of this novel relaxase family is available.

Table 2: General characteristics and properties of six main relaxase families summarized from Garcillán-Barcia *et al*. (110).

Family	Sub-families	Size	Structure	Conserved motifs	Active tyrosine	Special characteristics
MOB_F	2 MOB_{F1+F2}	very huge	$TrwC_{R388}$ $TraI_F$	I II III	1	large/complex *oriT* region; clusters with IS91 relaxases
MOB_Q	3 MOB_{Q1-Q3}	huge	$MobA_{RSF1010}$	I II III	1	clusters in part with MOB_P and MOB_{HEN} relaxases
MOB_P	7 MOB_{P1-P7}	largest family	-	I II III	1	well characterized; includes MOB_{HEN} family; clusters with MOB_Q
MOB_V	5 MOB_{V1-V5}	small	-	I III	no essential tyrosine identified	only few biochemical data available; possibly new mechanism; ancestrally related to MOB_P relaxases
MOB_C	3 MOB_{C1-C3}	small	-	no I +III equivalents (alternative DEE motif for metal binding)	sometimes substituted with Phe	unrelated to other relaxases; no covalent bond with DNA; mechanism unknown; also present in GIs; very few biochemical data
MOB_H	3 MOB_{H1-H3}	small	-	alternative 3H motif; conserved HD domains	rarely conserved tyrosin	very new family; present in plasmids, ICEs and GIs; unrelated to other relaxases; no biochemical data; mechanism unknown

To date, only little structural data of T4SS relaxases is available. The structures of the relaxase domains from $TraI_F$, $TrwC_{R388}$, and $MobA_{RSF1010}$ have been solved and shed light on the DNA binding and processing mechanism of these enzymes (227, 245, 246). In general all three structures share a

highly similar fold, even though the sequence identity between the different relaxases is relatively low (sequence identities: $MobA_{RSF1010}$ and $TraI_F$ → 12%; $MobA_{RSF1010}$ and $TrwC_{R388}$ → 10%; $TraI_F$ and $TrwC_{R388}$ → 42%) (232). The overall structure obtained from the relaxase domains of $TraI_F$, $TrwC_{R388}$ and $MobA_{RSF1010}$ has a hand like shape, formed by a 5-stranded anti-parallel β-sheet flanked by α-helices (232, 245). The active site is located within a binding cleft, which is formed by the anti-parallel β-sheet (232). The binding cleft contains three pockets with a diameter of 7-10 Å and a depth of 4-8 Å, which recognize certain bases with a "knob-into-hole"-like mechanism, enabeling the high sequence specificity of these enzymes (232). The size of the binding cleft is between 11-15 Å (232). The complete cleft is lined with neutral, nonaromatic residues resembling the ssDNA binding site (232, 245). The amino acid composition of the DNA binding site is a unique feature of relaxases and differs from other well-characterized ssDNA binding proteins (232). Only a few charged amino acids located in the binding cleft are sufficient to generate an electrostatic potential that positions with the help of a metal cofactor the scissile phosphate near the catalytic tyrosine (232). The metal binding site is located at the edge of the cleft in close proximity to the catalytic tyrosine (232). Relaxases are structurally and functionally related to virus replicases, rolling cycle transposases, replication initiator proteins and type IA topoisomerases (212, 247, 248).

The relaxase TraI of the novel T4SS of *N. gonorrhoeae* represents the prototype of the uncharacterized relaxase family MOB_H (110). Up to now, only little is known about the biochemical properties of TraI and its role in the DNA processing mechanism. TraI combines classical relaxase features with unique MOB_H characteristics that could not be found in the other relaxase families (110, 202). Like other relaxases contains the neisserial TraI an N-terminal relaxase domain, but does not share commonly conserved sequence motifs, except two conserved tyrosines (Y^{93} and Y^{201}) that have been reported to be important for DNA secretion (202). Instead of the classical metal binding site contains the neisserial TraI an altered histidine triad (H^{106} – H^{108}) as well as additional conserved HD domains, which are typically found in metal dependent phosphohydrolases (e.g. involved in the DNA metabolism) (202, 249). The HD domains of phosphohydrolases are involved in the coordination of metal ions, similar to the conserved histidine motif of relaxases (249). Both sequence motifs suggest a different metal binding mechanism within MOB_H relaxases. Mutagenesis studies and DNA secretion assays have shown that the conserved histidine triad of TraI does not have a significant influence on the level of DNA secretion (202). However, the mutation of the HD I domain (D^{120}) led to significantly reduced secretion levels (202). The relaxase domain is N-terminally flanked by a hydrophobic amphipathic α-helix that most likely acts as a membrane anchor to facilitate a close contact of TraI with the cytoplasmic membrane in close proximity to the secretion complex (202). This hydrophobic N-terminus is an uncommon and special feature of TraI and could not be found in other relaxases, even though different relaxases have been reported to attach to the cytoplasmic membrane (214). At its C-termius TraI contains a conserved DUF1528 domain, which is most likely responsible for protein-protein interactions (202). The precise function of this domain still remains unknown. Besides commonly shared protein characteristics also the genetic organization of the neisserial *dtr* genes corresponds to the typical organization of *dtr* genes in T4SSs. The neisserial *traI* gene is encoded together with *yaa*, *traD* and *yaf* in the same operon (the first operon of the GGI) and is in close proximity to the predicted *oriT* region (between *yaf* and *ltgX*) (202). Furthermore, initial evidence for a possible secretion of TraI has been reported by Salgado-Pabòn *et al*., based on the detection of 5' protected DNA in the external environment of the cells (202). It is not yet completely known, whether TraI is secreted to the outside of the cell or if the detected DNA results from external binding of TraI, which is transported via membrane blebs or delivered from autolysis.

1.9 Nicking accessory proteins

As a part of the relaxosome, nicking accessory proteins (NAPs) play an important role in the DNA processing mechanism of T4SSs. They are mostly rather small proteins with an average length of 200 amino acids (250). Genes encoding for NAPs are generally localized in close proximity to the relaxase and other *dtr* genes (219, 250, 251). In particular, NAPs fulfill several functions in the DNA processing mechanism comprising the transcriptional regulation of *dtr* genes (219, 252, 253), the spatial positioning of the relaxosome – DNA complex near the secretion channel (254-256), the recruitment of the relaxase (222, 254), and the enhancement of the endonucleic DNA processing activity of the relaxase (217-219). Considering the multi-tasking character of these proteins NAPs are of exceptional importance in regulating and specifying the DNA transfer within T4S. Well-studies examples of this protein group include e.g. $NikA_{R64}$, $TraJ_{RP4}$, $MobC_{RSF1010}$, $PcfF_{pCF10}$, $MbeC_{ColE1}$, $TraY_F$, $TraM_F$, $TrwA_{R388}$, $VirC2_{AT}$ and $VirC1_{AT}$ (215, 216, 218, 250, 254, 257-260).

The DNA binding of NAPs is often related to distinct binding sites, but is usually sequence unspecific. However, it has been reported that NAPs bind specifically dsDNA near the *oriT* sequence, in close proximity to the *nic* site and the binding sites of the IHF (218, 223, 259-262). A common ribbon-helix-helix domain structure (RHH) mediates the DNA binding and has been reported to be important for the recruitment of the relaxase (215, 250, 259, 260). The RHH binding fold of NAPs is structurally related to the phage p22 transcriptional regulators Arc and MetJ and to members of the CopG, Mnt and ParA family (215, 250, 263, 264). The sequence homology within different NAPs and RHH proteins is only poorly conserved with 15.4% similarity (265), which hampers the bioinformatical characterization and identification of these proteins. A conserved hydrophobic core in the RHH domain facilitates the DNA binding via polar amino acids localized in the β-sheet (260). Besides the conserved core of the RHH domain, only one conserved glycine can be identified connecting both α-helices (260).

Currently, two structures of NAPs are available, which are the crystal structure of $VirC2_{AT}$ at 1.7 Å resolution and the NMR structure of $NikA_{R64}$ (215, 260). Both resemble the RHH binding fold with two anti-parallel oriented β-sheets, which are characteristic for Arc/MetJ related RHH proteins. The anti-parallel β-sheets coordinate a group of polar amino acids that facilitate the DNA binding (260). The α-helices are most likely involved in the oligomerization of the proteins. The domain organization of several NAPs was analyzed with limited proteolysis and confirmed the presence of two different domains (250, 259, 260). The N-terminal domain represents the DNA binding domain, whereas the C-terminal domain is involved in protein-protein interactions mediating the tetramerization of DNA-bound proteins, as shown for $TrwA_{R388}$ and $MbeC_{ColE1}$ (250, 259). Actually, the C-terminal domain of $VirC2_{AT}$ is reported to be trypsin resistant, indicating the stable fold of this domain (260).

A common model of the functional mechanism of NAPs in T4SSs is based on the structural similarity to the Arc/MetJ transcription regulators and biochemical data obtained from several NAPs. Most of the NAPs bind DNA as homodimers and oligomerize to tetrameric states in their active conformation, similar to the mechanisms of Arc and MetJ (215, 250, 260). In solution most NAPs dimerize, with reported exceptions like $TraY_F$ which stays monomeric or $TrwA_{R388}$ forming stable tetramers (215, 250, 258-260). In general, the dsDNA is bound by two dimeric NAPs at specific binding sites within the *oriT* region (218, 259, 260). The conserved polar residues in the anti-parallel β-sheets facilitate a positively charged surface on one site of the dimer, enabeling the DNA binding (260). The close

proximity of the C-termini of both homodimers leads to a tetramerization via protein-protein interactions (259). The conformational changes during the tetramerization result in a bending of the dsDNA and enable an optimized binding and processing of the large relaxase at the *nic* site (215). Interactions of the NAPs with their cognate relaxase ensure the proper recruitment of the relaxase and the formation of the relaxosome complex (210, 218).

A well-studied example for the numerous protein-protein interactions of NAPs is the T4SS of *A. tumefaciens*. Biochemical data has shown that the spatial positioning protein $VirC1_{AT}$ binds to the "overdrive" sequence in the *oriT* region and interacts with $VirC2_{AT}$ forming a $VirC1_{AT}$/$VirC2_{AT}$ complex, which interact with the $VirD1_{AT}$/$VirD2_{AT}$ relaxase complex to recruit the relaxase to the *nic* site. The binding of $VirC1_{AT}$ to the "overdrive" sequence stimulates the DNA processing reaction and mediates the contact of the T-DNA with the $VirB/D4_{AT}$ secretion complex. The importance of the RHH domain for the DNA binding and the sequence specificity of the cleavage reaction has been studied in detail with DNA binding assays and mutagenesis studies for $VirC2_{AT}$ (102, 222, 260). Several other NAPs were studied with regard to their influence on the conjugative DNA transfer. The relaxosome of the *E. coli* F-plasmid is besides the Ti-plasmid of *A. tumefaciens* one of the best studied examples in conjugative DNA transfer. Both relaxosome complexes function in a similar mechanism. The F-plasmid relaxosome consists of several components necessary for a sufficient DNA processing, which are the relaxase $TraI_F$, the IHF_F, and $TraY_F$ (NAP), as well as the additional spatial positioning protein $TraM_F$ (210, 217, 218). $TraM_F$ exemplifies with its various protein interactions the importance of NAPs for the specificity of the DNA transfer reaction. The extensive interactions of $TraM_F$ with the coupling protein $TraD_F$ and the relaxase $TraI_F$ enable a highly specific DNA transfer during conjugation (256, 266). More detailed information on the interactions of $TraM_F$ with $TraD_F$ is given in Chapter 1.10. It has been reported that $TraM_F$ enhances the relaxation activity of $TraI_F$ up to 5-fold in presence of the other relaxosome components $TraI_F$, IHF_F, and $TraY_F$ (218).

The transcriptional regulation of *dtr* gene expression in T4SSs is a precisely controlled mechanism, which is important to allow a controlled conjugative DNA transfer. To ensure the sufficient processing and transfer of DNA during conjugation, the different relaxosome components need a distinct protein ratio for their proper interaction, which is often triggered by transcriptional regulation of the *dtr* operon, mediated by NAPs. For instance, $TraY_F$ binds to a sequence near its own promoter and up-regulates the transcription of $TraY_F$ (258). A down-regulation of the transcription of the complete *dtr* operon of the R388-plasmid encoded T4SS (*trwABC*) has been reported for $TrwA_{R388}$ (219). Both examples represent in part the versatile transcriptional regulation during DNA processing in T4SSs.

The DNA processing mechanism of the remarkable T4SS of *N. gonorrhoeae* is up to now unknown and needs to be studied in detail. Information about the relaxase TraI and the coupling protein TraD are given in Chapter 1.8 and 1.10. However, no further components of a putative relaxosome of *N. gonorrhoeae* were identified up to now. Since only little is known about the DNA processing mechanism of the T4SS of *N. gonorrhoeae*, it is most likely that additional relaxosome components have not been detected yet.

1.10 Coupling proteins

Coupling proteins (CPs) are essential for the function of T4SSs and are present in almost all known T4SSs, with one exception, the secretion system of the *B. pertussis* (230, 267, 268). CPs are of special importance in T4SSs since they mediate the interaction of the transferring substrate with the secretion complex (230). The sequence homology within different CPs is very low (15 – 20%) and only a few motifs are conserved (105). Based on conserved nucleotide binding domains and the number of transmembrane domains, up to six different classes of CPs can be distinguished (105). The main characteristics of these different classes are summarized in Table 3, according to the review of Alvarez-Martinez and Christie (105).

A well studied paradigm of Gram-negative coupling proteins is $TrwB_{R388}$. The crystal structure of the soluble C-terminus of $TrwB_{R388}$ shed light on possible mechanisms for the substrate translocation within T4SSs and was an important milestone within the research of T4SSs. The structural organization of $TrwB_{R388}$ revealed a strong similarity of $TrwB_{R388}$ to F_1F_0-ATPases, to the Cag525/HP0525 traffic ATPase, and to the T7 gene 4 DNA ring helicase (269, 270). The cytoplasmic C-terminal part of $TrwB_{R388}$ consist of two different domains, a nucleotide binding domain (NBD) and an all-α domain (AAD), whereas the N-terminal part contains the transmembrane domains (TM) and is integrated into the inner membrane (269, 270). The C-terminus of CPs is unstructured and contains the AAD, which shows similarity to the XerD recombinase and is involved in the binding of DNA substrates and protein-protein interactions (269, 270). The NBD is localized between the AAD and the TM domain and is structurally related to RecA and the T7 gene 4 ring helicase (269). The Walker A and B motifs for ATP binding and hydrolysis are the most conserved motifs within CPs and are both localized in the NBD (105). Mutations in the Walker motifs lead to conjugation deficient phenotypes and demonstrate the importance of ATP hydrolysis and the NBD for the proper function of CPs (271-273). The TM domains are responsible for the integration of the CPs into the cytoplasmic membrane, similar to F_1F_0-ATPase (270, 274-276). The holoenzyme assembles as a homohexamer in presence of the relaxosome complex (277, 278). The assembled hexameric CP is located in close proximity to the secretion complex and is supposed to energize the translocation process by ATP hydrolysis (230, 272, 273). The homohexameric structure of TrwB has an orange like shape with a diameter of 110 Å and a central channel in the middle with 20 Å in diameter. The central channel is constricted to 7-8 Å, which is supposed to prevent a leakage of the cytoplasmic membrane and could most likely be opened and closed by conformational changes e.g. during ATP hydrolysis. The height of the hexameric complex is 90 Å (270).

The interaction of CPs with processed DNA substrates and proteins of the relaxosome complex could be shown *in vitro* for several CPs. In most cases the initial substrate interaction with the CP starts via a protein-protein interaction at the C-terminus of the CP and only in a few exceptions the CP interacts directly with the processed DNA. Protein-protein interactions with CPs have been shown for different relaxases and nicking accessory proteins. For instance, $VirD4_{AT}$ needs for the binding of ssDNA the enzymatic activity of the relaxase $VirD2_{AT}$ (137). The interaction of CPs with their cognate relaxases could be also shown for Gram-positive T4SSs, exemplified with the interactions of $PcfC_{pCF10}$ and the relaxase $PcfG_{pCF10}$ and $Orf10_{pIP501}$ and TraA (254, 279). Intensive interactions of the hexameric CP with nicking accessory proteins of the relaxosome are often the key for the highly specific substrate transfer in T4SSs. This has been shown for $TraD_F$, which interacts extensively with the nicking accessory protein $TraM_F$ (220, 256, 266). Interestingly, it has been shown that CPs can

substitute each other within different T4SSs, but the interchangeability seems to be restricted to the different classes of CPs (105). It has been reported that $TraG_{RP4}$ can substitute $TrwB_{R388}$, and $TraG_{pTic58}$ can be exchanged with $VirD4_{AT}$ (268, 280). The compatibility of CPs from different T4SSs among each other indicates the possibility of different ancestral origins of the *dtr* proteins and the *mpf* proteins, which is in accordance to bioinformatical data discussed in the context of integrative and conjugative elements in Chapter 1.5.

Table 3: Overview on the classification of CPs summarized from Alvarez-Martinez and Christie, 2009 (105).

Class	Characteristics	Examples
VirD4 like CPs	well characterized class of CPs at least 2 TM-domains structure of $TrwB_{R388}$ available huge sequence heterogeneity large proteins with 600 – 750 AS	$VirD4_{AT}$; $TrwB_{R388}$ (2 TM-domains) $YdcQ_{ICEBs1}$; $Orf14_{pSKU146}$ (2 TM-domains) $HP0524_{Hyp}$; $TraG_{pSymA}$ (4 TM-domains) $TraG_{B9}$ (4 TM-domains) $TraD_F$; $TraG_{RIc}$ (2 TM-domains; C-terminal extension) $TraD_{NG}$ (no TM-domain; C-terminal extension)
FtsK/SpoIIIE like CPs	variability in length differ in TM-domain number function independent from other T4SS components	$TraG_{pKLC102}$; $Orf16_{pAB559}$ (3 TM-domains) $TraB_{pSGs}$ (1 TM-domain) $TcpA_{pCW3}$ (2 TM-domains; C-terminal extension)
$TraG_{R27}$ + TraJ like CPs	short TM-domains need additional TraJ (with 4 TM-domains)	$TraG_{R27}$; $TraG_{R478}$; $p081_{pWWO}$
CPs without TM-domain + small membrane protein	no TM-domains linked with small predicted membrane protein (2TM-domains)	$TraG_{BC}$; $Orf10_{pIP501}$; $pPSR1p49_{pPSR1}$
CPs without TM-domain	-	$BACCAP01795_{Bca}$; $CLOL250\_01204_{Clo}$
Archaeal CPs	archaeal specific sequences no TM-domains	$TraG_{pKEF9}$; $Orf1023_{pAH1}$

Currently, three different models for the DNA translocation mechanism mediated by CPs have been proposed. One model postulates the CP functioning as a connector between the secretion complex and the relaxosome-DNA complex. The recognition of the secretion substrate and the secretion complex on both sites of the CP does not require ATP hydrolysis, but it is assumed that conformational changes due to ATP hydrolysis could act as a molecular switch in the CP to trigger the secretion process (230). Both of the other discussed models require for their postulated translocation mechanisms rotational force generated by ATP hydrolysis and are due to the ATPase activity of CPs commonly the more preferred models. One model proposes the CP acting as a DNA pumping motor with the processed ssDNA wrapped around the cytosolic part of the protein. The binding of ssDNA and dsDNA is enabled by a positively charged belt at the external surface of the CP.

The hydrolysis of ATP leads to the rotation of the CP and triggers the bound relaxosome complex to the secretion pore (230). Since only very few biochemical data supports this model it is one of the less preferred mechanisms. The up to now most preferred model for the DNA translocation via CPs is similar to the mechanism of SpoIIIE and FtsK. In this model the processed ssDNA substrates pass through the central channel of the CP, which is acting like a DNA pump. The inner surface of the channel is negatively charged, which would at first instance hinder the translocation of negatively charged DNA, but similar to the external surface of the CPs a positively charged belt enables the translocation of the DNA. ATP hydrolysis will force the pumping of the DNA and it is supposed that conformational changes in CPs trigger the opening and closing of the central channel (230).

The CP TraD of *N. gonorrhoeae* is essential for the DNA secretion (198), but so far no biochemical data of this protein is available. Interestingly, TraD is classified to the VirD4 like CPs (105), but seems to lack the TM domains typically found within this class, and is in one operon with the small membrane protein Yaa that contains at least four predicted TM domains. This clustering may result from the classification based on the conserved Walker A and B motifs, but considering the predicted TM domain topology and the genetically close proximity to *yaa* it is more likely that the neisserial TraD clusters into the class of $TraG_{R27}$ like CPs or the class of TM domain-less CPs associated with small predicted membrane proteins. A phylogenetic analysis of different CPs performed by E. Pachulec, a former PhD student of our group, revealed that the neisserial TraD clusters not as expected to $TraD_F$, but seems to be closer related to $TraG_{R27}$, $TraD_{R391}$ and $TraD_{SXT}$ (PhD thesis E. Pachulec, 2010). This strengthens the hypothesis that TraD belongs to the class of $TraG_{R27}$ like CPs and the function of Yaa is may be similar to that of TraJ. However, the biochemical characterization of TraD is of oustanding importance to shed light on the DNA translocation mechanism of the novel T4SS of *N. gonorrhoeae*, which is in the focus of this study.

2. Objectives of the study

The current study focuses on the understanding of the DNA secretion mechanism of the remarkable T4SS of *N. gonorrhoeae*, which is the only known T4SS that secretes DNA directly into the environment, without a previous cell-cell contact. This novel T4SS encodes an *mpf* complex which is highly similar to the T4SS of the *E. coli* F-plasmid, but the *dtr* genes seem to be relatated to another system since phylogenetic studies with related T4SSs revealed a different clustering of the neisserial and the *E. coli dtr* genes.

The key enzyme for the DNA processing in T4SSs is the relaxase, which introduces a nick into the dsDNA and unwinds the DNA to generate a single stranded transfer DNA. The neisserial relaxase TraI represents a prototype of the novel relaxase family MOB_H and up to now, no biochemical data of this relaxase family is available. Since TraI contains several sequence motifs (e.g. a hydrophobic N-terminus, conserved HD domains, an altered histidine motif, and a conserved C-terminal DUF1528 domain) that could not be found in other relaxases, it is assumed that TraI (representative for MOB_H relaxases) has most likely a different enzymatic mechanism. To shed light on the unknown DNA processing mechanism of MOB_H relaxases and to understand the mechanism of DNA transfer within the novel T4SS of *N. gonorrhoeae*, it is of outstanding importance to characterize the biochemical properties of the key enzyme TraI, which is in the major focus of the current study.

In T4SSs, the contact of the processed transfer DNA with the secretion complex is mediated via a coupling protein (TraD), which is an integral inner membrane ATPase localized in close proximity to the secretion channel. The mechanism of how coupling proteins transfer the ssDNA to the secretion channel is still under discussion. To elucidate the DNA transfer mechanism of the remarkable T4SS of *N. gonorrhoeae*, it is necessary to understand the enzymatic activity and functional properties of the neisserial coupling protein TraD. Therefore, the biochemical characterization of TraD is besides the characterization of the relaxase TraI another aspect of the current study.

The DNA cleavage reaction of the relaxase is often initiated through the presence of additional nicking accessory proteins that form together with the relaxase and the DNA the “relaxosome” complex. Up to now, no additional relaxosome components have been identified in the T4SS of *N. gonorrhoeae*. However, the possibility of the formation of a relaxosome in *N. gonorrhoeae* cannot be absolutely excluded. One subject within the current study is the identification of a possible nicking accessory protein within the T4SS of *N. gonorrhoeae* that is involved in the DNA processing mechanism and possibly in the formation of a neisserial relaxosome. In addition to the relaxase gene the *dtr* operon contains two hypothetical genes (*yaa* and *yaf*) encoding for proteins with up to now unknown function. Considering bioinformatical evidences, Yaf seems to be a promising candidate for a nicking accessory protein, and the functional and biochemical characterization of this hypothetical protein in context with a possible relaxosome formation in *N. gonorrhoeae* is within the focus of the current study.

The current study should provide information about the enzymatic function of three up to now uncharacterized key proteins in the DNA processing mechanism of *N. gonorrhoeae*, and should improve the understanding of the novel DNA secretion mechanism of the neisserial T4SS.

3. Material and Methods

3.1 Material

3.1.1 Chemicals and enzymes

All chemicals were purchased from Carl Roth, Sigma-Aldrich, Merck, Jena Bioscience and Hartmann Analytics. Enzymes were purchased from New England Biolabs, Roche or Fermentas.

3.1.2 Media

LB medium

LB media (Luria/Miller) and LB agar (Luria/Miller) were purchased from Roth. The composition of LB medium and LB agar is given in Table 4 and 5. The media were autoclaved for 20 min at 120°C.

Table 4: Components for liquid LB medium

Concentration	Component
10 g/L	Tryptone
5 g/L	Yeast extract
10 g/L	NaCl

Table 5: Components for LB agar medium

Concentration	Component
10 g/L	Tryptone
5 g/L	Yeast extract
10 g/L	NaCl
15 g/L	Agar-Agar

SOC medium

SOC medium was used for the initial growth of transformed *E. coli* cells. The composition for SOC medium is given in Table 6. All components except glucose were mixed and autoclaved. A 50% (w/v) glucose solution was autoclaved separately and was added to the medium to its final concentration afterwards.

Table 6: Components for SOC medium

Concentration	Component
2% (w/v)	Tryptone
0.5% (w/v)	Yeast extract
8.6 mM	NaCl
10 mM	$MgCl_2$
10 mM	$MgSO_4$
2.5 mM	KCl
➢ filled up to 1 L and autoclaved	
0.36% (v/v)	Glucose

Autoinduction medium

The method of autoinduction was used for the overproduction of full-length TraD and N-terminal truncated TraD. The composition of the medium was in accordance to Studier *et al.* (281). The sugar and salt solutions, as well as the basic media components, were autoclaved separately from each other. The trace element solution was filtered with a 0.22 µm sterile syringe filter. The components were mixed under sterile conditions shortly before the inoculation with *E.coli* cells.

3.1.3 Buffer and solution list

Isolation of TraD

Denaturing buffer		***Purification buffer***	
50 mM	HEPES/NaOH pH 7.5	50 mM	HEPES/NaOH pH 7.5
6 M	Guanidine/HCl	6 M	Guanidine/HCl
		1% (v/v)	TritonX-100

Isolation of MBP-TraI fusionprotein

His-tag MBP-TraI A		***His-tag MBP-TraI B***	
30 mM	Tris/HCl pH 8.0	30 mM	Tris/HCl pH 8
150 mM	NaCl	150 mM	NaCl
5 mM	Imidazole	400 mM	Imidazole
5% (v/v)	Glycerol	5% (v/v)	Glycerol

Gelfiltration MBP-TraI	
30 mM	Tris/HCl pH 8.0
150 mM	NaCl
5% (v/v)	Glycerol

Isolation of native TraI without detergent

Q Sepharose TraI A		***Q sepharose TraI B***	
50 mM	Tris/HCl pH 7.6	50 mM	Tris/HCl pH 7.6
		1 M	NaCl

Phenyl-Sepharose TraI A		***Phenyl-Sepharose TraI B***	
50 mM	Tris/HCl pH 7.6	50 mM	Tris/HCl pH 7.6
40 mM	$(NH_4)_2SO_4$	2 M	$(NH_4)_2SO_4$

Buffer gelfiltration native TraI Phen

50 mM	Tris/HCl pH 7.6
150 mM	NaCl

Isolation of native TraI with detergent

Ion exchange TraI detergent A		***Ion exchange TraI detergent B***	
50 mM	Tris/HCl pH 7.6	50 mM	Tris/HCl pH 7.6
25 mM	NaCl	1 M	NaCl
10 % (v/v)	Glycerol	10 % (v/v)	Glycerol
0.05% (v/v)	DDM	0.05% (v/v)	DDM

Gelfiltration TraI detergent

50 mM	Tris/HCl pH 7.6
150 mM	NaCl
10% (v/v)	Glycerol
0.05% (v/v)	DDM

Isolation of His-tagged yaf

His-tag Yaf A		***His-tag Yaf B***	
50 mM	HEPES/NaOH pH 7.0	50 mM	HEPES/NaOH pH 7.0
500 mM	NaCl	500 mM	NaCl
10 mM	Imidazole	400 mM	Imidazole
10% (v/v)	Glycerol	10% (v/v)	Glycerol

Gelfiltration Yaf (His-tag)

50 mM	HEPES/NaOH pH 7.0
250 mM	NaCl
10% (v/v)	Glycerol

Isolation of native Yaf

Ion exchange Yaf A

50 mM	HEPES/NaOH pH 7.0
50 mM	NaCl
10% (v/v)	Glycerol

Ion exchange Yaf B

50 mM	HEPES/NaOH pH 7.0
1 M	NaCl
10% (v/v)	Glycerol

Gelfiltration native Yaf

50 mM	HEPES/NaOH pH 7.0
150 mM	NaCl
10% (v/v)	Glycerol

Electrophoresis buffer

10x TB buffer

108 g/L	Tris base
55 g/L	Boric acid

- adjust to pH 8.0 with HCl

10x TGS buffer

30.3 g/L	Tris base
144 g/L	Glycine
10 g/L	SDS

Stacking buffer (SDS-PAGE)

0.5 M	Tris base
0.4% (w/v)	SDS

- adjust to pH 6.8

Separation buffer (SDS-PAGE)

1.5 M	Tris base
0.4% (w/v)	SDS

- adjust to pH 8.8

Cathode buffer A (BN)

50 mM	Tricine
7.5 mM	Imidazole
0.02% (w/v)	SERVA Blue G

Cathode buffer B (BN)

50 mM	Tricine
7.5 mM	Imidazole
0.002% (w/v)	SERVA Blue G

Anode buffer (BN)

25 mM	Imidazole/HCl

- adjust to pH 7.0

BN Gelbuffer 3x

75 mM	Imidazole/HCl pH 7.0
1.5 M	6-aminocaproic acid

5x SDS-PAGE loading dye

10% (w/v)	SDS
50% (v/v)	Glycerol
0.04%(w/v)	Bromophenol blue
500 mM	DTT
300 mM	Tris/HCl pH 6.8

Coomassie staining solution

0.1% (w/v)	Coomassie R
0.1% (w/v)	$CuSO_4$
25% (v/v)	Ethanol
20% (v/v)	Acetic acid

Destaining solution

0.5% (w/v)	$CuSO_4$
7% (v/v)	Acetic acid
12% (v/v)	2-Propanol

DNA processing assays

6x DNA loading dye		***5x Colorless loading dye***	
50% (w/v)	Sucrose	150 mM	Tris/HCl pH 7.5
0.1% (w/v)	Bromophenol blue	60% (v/v)	Glycerol

2x Stop solution		***5x Native loading dye***	
80% (v/v)	Formamide	150 mM	Tris/HCl pH 7.5
1% (v/v)	SDS	60% (v/v)	Glycerol
40 mM	$Na_2EDTA * 2 H_2O$	0.1% (w/v)	Bromophenol blue

10x TE buffer

100 mM Tris base

10 mM $Na_2EDTA * 2 H_2O$

- adjust to pH 8.0 with HCl

3.2 Strains, oligonucleotides and plasmids

3.2.1 Strains

Table 7: Strains used for cloning and protein overexpression

Strain	Genotype	Reference	Source
E. coli DH5α	*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ) M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	(282)	Invitrogen
E. coli BL21 (DE3)	F^- *ompT* $hsdS_B(r_B^- m_B^-)$ *gal dcm* (DE3)	(283)	NEB
E. coli Tuner (DE3)	F^- *ompT* $hsdS_B$ ($r_B^- m_B^-$) *gal dcm lacY1*(DE3)	(284)	Novagen
E. coli C43 (DE3) pRARE	F^- *ompT gal dcm* $hsdS_B(r_B^- m_B^-)$*(DE3) pRARE* (Cam^R)	(285)	Lucigen
E. coli Arctic Express (DE3)	*E. coli B F– ompT hsdS* (rB^-mB^-) dcm^+ *Tetr gal λ(DE3) endA Hte [cpn10 cpn60 Gentr]*	(286) (287)	Stratagene
E. coli LEMO21 (DE3)	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS/ pLemo((CamR)(pACYC184-PrhaBAD-lysY))*	(288)	NEB
E. coli BL21 (DE3) Star	F^- *ompT* $hsdS_B$ (r_B^-, m_B^-) *gal dcm rne131 (DE3)*	(289) (290)	Invitrogen

3.2.2 Oligonucleotides

Table 8: Oligonucleotides used for cloning and mutagenesis. Oligonucleotides are shown from 5' to 3'.

Name	Sequence	Description
native *traI* F	ACTAGTAAAAATCGGCATCATGAAAACAAGCCTTCTCAC	SpeI site; start
native *traI* R	TCAGGGAAGCGGCCGCTCATTTTTGTTCCATTACTAATAAG	NotI site; end
eXact-tag *traI* F	GCGCGGACTAGTATGAAAACAAGCCTTCTCAC	SpeI site; start
native *yaf* F	GGCCGCGCATATGAATGAAACAGATCTTCCCAAGCTTCTTAGTG	NdeI site; start
native *yaf* R	GCGGCGCTCGAGTCATGATGATTTAAGCAAAGATTTTAG	XhoI site; end
traI Y93F F	GCCCTTCCCTTTTTGGAAAAATTCATAGCC	Y93F mutation
traI Y93F R	CAAAGGCTATGAATTTTTCCAAAAAGG	Y93F mutation
MBP-His-tag XhoI	CGCATCTCGAGTGGTCCTTGTTGGTGAAGTG	N-term. His-tag; XhoI site
MBP-TEV XbaI	GCGGCTCTAGATCCTTGGAAGTACAGGTTCTCGGATCCGAATTCCGGTACC TTG	C-term. TEV site; XbaI site
Vector XhoI	CGCGGCTCGAGATGCTATGGTCCTTGTTGGTGAAGTG	Vectoramplification; XhoI site
Vector XbaI	CGGAATTCGGATCCTCTAGAGTC	Vectoramplification; XbaI site
XhoI removal F	GGACCATAGCATATGCACCACCATCATCACCATC	Removes XhoI site
XhoI removal R	GATGGTGGTGCATATGCTATGGTCCTTGTTGGTGAAGTG	Removes XhoI site
His-tag *yaf* F	GACGCGCATATGCACCATCACCACCATCATCACCATCACCACAATGAAACA GATCTTCCCAAGCTTCTTAGTG	NdeI site; start
MBP-TEV-TraI F	CGCAGTCTAGAATGAAAACAAGCCTTCTCAC	XbaI site; start
MBP-TEV-TraI R	GCGATCCTGCAGGTCATTTTTGTTCCATTACTAATAAG	SbfI site; end
yaf Y47F F	CGTCATCATTTCCAGAAAACCGCTAGCACGCTGTATCC	Y47F mutation
yaf Y47F R	GCGTGCTAGCGGTTTTCTGGAAATGATGACGCTGGATGAC	Y47F mutation
yaf Y100F/Y104F F	CTTCAAGTGAATTCCAGGCTTCAGAACCGGCTTGGGCCATTAG	Y100F/Y104F mut
yaf Y100F/Y104F R	ATCTTATTTCTAATGGCCCAAGCCGGTTCTGAAGCCTGGAATTCAC	Y100F/Y104F mut
yaf D89A F	TTTGCAGAAGCACTTGCCGCTGACTTTGAAATC	D89A mutation
yaf D89A R	GATTTCAAAGTCAGCGGCAAGTGCTTCTGCAAAAGTCG	D89A mutation
yaf D90A F	GCAGAAGCACTTGCCGATGCCTTTGAAATCTTATTTCTAATGG	D90A mutation
yaf D90A R	GATTTCAAAGGCATCGGCAAGTGCTTCTGC	D90A mutation
yaf E57A F	GGTTGCCGTATTAGCATCGTCATCCAGCGTC	E57A mutation
yaf E57A R	CTGGATGACGATGCTAATACGGCAACCTTC	E57A mutation
yaf F91A F	GATTTCAGCGTCATCGGCAAGTGCTTCTG	F91A mutation
yaf F91A R	CTTGCCGATGACGCTGAAATCTTATTTCTAATGG	F91A mutation
yaf F115A F	GTCTAATACAAGCTTCCTCAAAGCTTTCTTCAAG	F115A mutation
yaf F115A R	GAGGAAGCTTGTATTAGACATACTACGCTACC	F115A mutation
yaf S107A F	GCTTTCTTCAAGAGCATTCCAGTATTCAGAACCG	S107A mutation
yaf S107A R	GAATACTGGAATGCTCTTGAAGAAAGCTTTGAGG	S107A mutation

Table 9: Oligonucleotides used for enzymatic activity assays

Name	Sequence	Description
GGI oriT 75bp F	AATGCTGCCCCTTTTAATAGGCAAAGGCCTGCATTTTTACATGCAGGCCTTTGTTTTACCTTTAACTTTTACCCT	ssDNA 75 bp long; contain GGI oriT
GGI oriT 75bp R	AGGGTAAAAGTTAAAGGTAAAACAAAGGCCTGCATGTAAAAATGCAGGCCTTTGCCTATTAAAAGGGGCAGCATT	ssDNA 75 bp long; contain GGI oriT
GGI oriT 120bp F	CAGAATAAGATTTATAAGGTATAATGCTGCCCCTTTTAATAGGCAAAGGCCTGCATTTTTACATGCAGGCCTTTGTTTTACCTTTAACTTTTACCCTTTTTCCCTTGAACCCTTCCTTTA	ssDNA 120 bp long; GGI oriT
GGI oriT 120bp R	TAAAGGAAGGGTTCAAGGGAAAAAGGGTAAAAGTTAAAGGTAAAACAAAGGCCTGCATGTAAAAATGCAGGCCTTTGCCTATTAAAAGGGGCAGCATTATACCTTATAAATCTTATTCTG	ssDNA 120 bp long; GGI oriT
F oriT 76bp F	GTTCTCACCACCAAAAGCACCACACCCCACGCAAAAACAAGTTTTTGCTGATTTTTCTTTATAAATAGAGTGTTAT	ssDNA 76 bp long; contain F oriT
F oriT 76bp R	ATAACACTCTATTTATAAAGAAAAATCAGCAAAAACTTGTTTTTGCGTGGGGTGTGGTGCTTTTGGTGGTGAGAAC	ssDNA 76 bp long; contain F oriT
F oriT 120bp F	CTCAACAGGTTGGTGGTTCTCACCACCAAAAGCACCACACCCCACGCAAAAACAAGTTTTTGCTGATTTTTCTTTATAAATAGAGTGTTATGAAAAATTAGTTTCTCTTACTCTCTTTAT	ssDNA 120 bp long; F oriT
F oriT 120bp R	ATAAAGAGAGTAAGAGAAACTAATTTTTCATAACACTCTATTTATAAAGAAAAATCAGCAAAAACTTGTTTTTGCGTGGGGTGTGGTGCTTTTGGTGGTGAGAACCACCAACCTGTTGAG	ssDNA 120 bp long; F oriT
GGI oriT F	GAAGCTTGGGAAGATCTG	used for GGI *oriT* PCR product
GGI oriT R	GGGTAAAGGCTCTTATCC	used for GGI *oriT* PCR product (short)
GGI oriT ltgX R	ATGGCCCTTAATCTCATGTC	used for GGI *oriT* PCR product (long)
914	ACCGGCTACACGTACAACTG	used for *dif*-site PCR product
yaa-sequence	GAGTGCAGGCATTTATCACC	used for *dif*-site PCR product

3.2.3 Plasmids

Table 10: Plasmids used for protein overproduction and enzymatic activity assays

Plasmid	Vector	Gene	Res.	Source	Oligonucleotide	Method
pMW001	pET15b	TraD (C-teminal His-tag)	Amp	M. Wiezes (AG Does)		
pMW002	pET15b	TraD (N-terminal His-tag)	Amp	M. Wiezes (AG Does)		
pMW003	pET15b	TraD (Δ TM; C-term. His-tag)	Amp	M. Wiezes (AG Does)		
pMW004	pET15b	TraD (Δ TM; N-term. His-tag)	Amp	M. Wiezes (AG Does)		
pEP073	pCDFDuet	390 bp *oriT* of F-plasmid	Strep	E. Pachulec (AG Does)		
pEP074	pCDFDuet	468 bp *oriT* of GGI	Strep	E. Pachulec (AG Does)		
-pJK001	pPAL7	*traI* (native)	Amp	E.-M. Heller J. Koch	native *traI* F: native *traI* R	cloning with ligation
pJK002	pPAL7	*traI* (N-terminal eXact-tag)	Amp	E.-M. Heller J. Koch	eXact-tag *traI* F; native *traI* R	cloning with ligation
pKS008	pET20b(+)	*yaf* (native)	Amp	E.-M. Heller K. Siewering	native *yaf* F; native *yaf* R	cloning with ligation
pSJ007	pKH184	*traI* (N-terminal His-tag)	Kan	S. Jain (AG Does)		
pSJ011	pKH184	*traI* (C-terminal His-tag)	Kan	S. Jain (AG Does)		
pSJ017	IBA-101	*traI* (C-terminal Strep-tag)	Amp	S. Jain (AG Does)		
pSJ018	IBA-102	*traI* (N-terminal Strep-tag)	Amp	S. Jain (AG Does)		
pSJ024	pTE088	*traI* (N-term. MBP fusion)	Amp	S. Jain (AG Does)		
pEMH004	pJK001	*traI* (native); Y93F mutant	Amp	E.-M. Heller	*traI* Y93F R; *traI* Y93F F	mutagenesis
pEMH007	pMALc2E	introduced TEV cleavage site	Amp	E.-M. Heller	MBP-His-tag XhoI; MBP-TEV XbaI; Vector XhoI; Vector XbaI; XhoI rem. F; XhoI rem. R	cloning with ligation and mutagenesis
pEMH008	pET20b(+)	*yaf* (N-terminal His-tag)	Amp	E.-M. Heller	His-tag *yaf* F; native *yaf* R	cloning with ligation
pEMH010	pEMH007	*traI* (N- term. His-MBP-TEV fusion)	Amp	E.-M. Heller	MBP-TEV-TraI F; MBP-TEV-TraI R	cloning with ligation
pEMH017	pEMH010	*traI* (N-MBP fusion); Y93F mutant	Amp	E.-M. Heller	*traI* Y93F R: *traI* Y93F F	mutagenesis
pEMH021	pEMH008	*yaf* (N-His-tag); Y47F mutant	Amp	E.-M. Heller	*yaf* Y47F R; *yaf* Y47F F	mutagenesis
pEMH022	pEMH008	*yaf* (N-His-tag); Y100F/Y104F mut.	Amp	E.-M. Heller	*yaf* Y100F/Y104F F; *yaf* Y100F/Y104F R	mutagenesis
pEMH023	pEMH008	*yaf* (N-His-tag); D89A mutant	Amp	E.-M. Heller	*yaf* D89A F; *yaf* D89A R	mutagenesis
pEMH024	pEMH008	*yaf* (N-His-tag); D90A mutant	Amp	E.-M. Heller	*yaf* D90A F; *yaf* D90A R	mutagenesis
pEMH025	pEMH008	*yaf* (N-His-tag); D89A/D90A mut.	Amp	E.-M. Heller	*yaf* D89A F; *yaf* D89A R *yaf* D90A F; *yaf* D90A R	double mutagenesis
pEMH038	pEMH008	*yaf* (N-His-tag); E57A mutant	Amp	E.-M. Heller	*yaf* E57A F; *yaf* E57A R	mutagenesis
pEMH041	pEMH008	*yaf* (N-His-tag); F91A mutant	Amp	E.-M. Heller	*yaf* F91A F; *yaf* F91A R	mutagenesis
pEMH043	pEMH008	*yaf* (N-His-tag); F115A mutant	Amp	E.-M. Heller	*yaf* F115A F; *yaf* F115A R	mutagenesis
pEMH045	pEMH008	*yaf* (N-His-tag); S107A mutant	Amp	E.-M. Heller	*yaf* S107A F; *yaf* S107A R	mutagenesis

3.3 Cloning and mutagenesis

3.3.1 PCR for cloning

The genes of interest were fused 5' and 3' with recognition sites for the corresponding restriction enzymes and were amplified with PCR. The restriction enzyme recognition sites were introduced with the corresponding primer pair and are given in Table 8. The PCRs were performed with Phusion polymerase (Finnzymes) as described in Table 11 and 12. In case of the amplification of neisserial genes *N. gonorrhoeae* MS11 chromosomal DNA was used as template DNA. The annealing temperature varied corresponding to the used primer pairs and was calculated for each PCR with the Clone Manager Professional 9.0 (Sci-ed Software). The elongation time of the product was changed corresponding the length of the gene of interst and was calculated with 30 sec/kb. The PCR products were purified with the PCR clean up kit (Zymo) as described in the manual and were analyzed on agarose gels. The concentration of the PCR product was measured with the Nanodrop.

Table 11: PCR reaction-mix for cloning

Volume	Components
10 µl	10x HF-buffer (or GC-buffer)
1 µl	dNTP-mix (10 µM)
0.25 µl	Oligonucleotide 1 (0.5 µM)
0.25 µl	Oligonucleotide 2 (0.5 µM)
0.75 µl	Template DNA (< 250 ng)
37.5 µl	deionized autocalved H_2O
0.25 µl	Phusion

Table 12: PCR program for cloning

Temperature	Time	Cycles
98°C	2 min	1x
98°C	10 sec	35x
X°C	30 sec	
72°C	30 sec/kb	
72°C	10 min	1x
4°C	∞	-

3.3.2 Restriction digests for cloning

The purified PCR products were mixed with the corresponding restriction enzymes and were digested as recommended by the manufacturer. In most cases the reaction mix was incubated in a double digest reaction for 1 h at 37°C. The restriction enzymes were inactivated as recommended by the manufacturer and the digested DNA was purified with the Clean & Concentrator kit (Zymo) as described in the manual. The DNA concentration was measured with Nanodrop.

Table 13: Reaction-mix for restriction digests

Volume	Component
25 µl	Purified PCR product
5 µl	10x reaction buffer (NEB1-4)
0.5 µl	Restriction enzyme 1 and 2
0.5 µl	100x BSA (optional)
18.5 µl	deionized autoclaved H_2O

3.3.3 Ligation of DNA fragments

The vectors were digested as described above and were additionally incubated with alkaline phosphatase as recommended by the manufacturer, to avoid religation of the original vector. The amounts of insert and vector were calculated as following:

Insert mass (ng) = 6* ((Insert length (bp) / Vector length (bp)) * Vector mass (ng))

The ligation reaction was prepared as described in Table 14 and was incubated for 1 h at 22°C. 15 µl of the reaction were transformed into chemical competent *E. coli* DH5α cells. The transformed cells were plated on LB plates with the corresponding antibiotic and single colonies were picked and used for the isolation of the generated construct.

Table 14: Reaction-mix for ligation reaction

Volume	Component
X µl	Insert gene
X µl	Vector
4 µl	5x rapid ligation buffer
1 µl	T4 ligase
X µl	deionized autoclaved H_2O
➢ total reaction volume: 20 µl	

3.3.4 Transformation of *E. coli*

Unless otherwise noted, chemical competent *E. coli* cells were transformed as described following. 200 µl of frozen competent cells were thawed on ice and 0.25 µl of plasmid DNA was added. The cells were incubated together with the DNA for 5 min on ice prior a 45 sec heat-shock at 42°C. Afterwards the cells were incubated on ice for 10 min before the addition of 800 µl LB medium. Subsequently the cells were incubated in a shaking incubator for 1 h at 37°C and were either plated on antibiotic containing LB plates or were directly inoculated with 50 ml LB medium with antibiotic in case of pre-cultures for the overproduction of the protein.

3.3.5 Amplification and isolation of plasmids

Plasmids were transformed into chemical competent *E. coli* DH5α cells as described in Chapter 3.3.4. A 50 ml over night culture with the corresponding antibiotic was incubated in a shaking incubator at 37°C. 3-5 ml of the overnight culture was centrifuged (Eppendorf *Fresco* Table centrifuge; 3.500 rmp; 4°C; 4 min) and the pellet was used for plasmid isolation. Plasmids were isolated with the *Classic* Miniprep Kit (Zymo) as described in the manual. The DNA concentration was determined with Nanodrop (Peqlab) and the plasmids were stored at -20°C until further use.

Plasmids used in DNA relaxtion assays described in Chapter 3.8.6 were isolated with the Midiprep-Kit (Quiagen) as described in the manual. DNA coiling was analyzed with agarose gel electrophoresis as described in Chapter 3.7.2. The concentration of the isolated plasmids was determined with Nanodrop (Peqlab) and the plasmids were stored at 4°C.

The cloned and mutagenized constructs were analyzed for theire correct sequence by sequencing the distinct gene of interest. The sequencing reaction was performed by the company "Eurofins MWG Operon". The data was analyzed with Clone Manager Professional 9.0 (Sci-ed Software).

3.3.6 Site-directed mutagenesis

Site-directed mutagenesis was performed to generate point mutations in accordance with the protocol of the QuickChange site-directed mutagenesis kit from Stratagene. A list of used oligonucleotides is given in Chapter 3.2.2. The PCR-program and the reaction-mix are given in Table 15 and 16. Yaf mutants were generated using plasmid pEMH008 (WT-Yaf with N-terminal 10x His-tag) as template, whereas TraI mutants were generated using pEMH010 (MBP-TraI fusion) and pJK001 (native TraI) as templates. The PCR products were digested with DpnI for at least for 4 h at 37°C. The reaction mix for the DpnI digest is given in Table 17. The digested PCR product was purified with the DNA Clean & Concentrator Kit 25 (Zymo) following the manufacturer's protocol. 5-10 µl of the purified PCR products were transformed as described in Chapter 3.3.4 into chemical competent *E. coli* DH5α cells. Single colonies were picked from plates and plasmids were amplified as described in Chapter 3.3.5.

Table 15: PCR reaction-mix for site-directed mutagenesis

Volume	Component
10 µl	5x KapaHIFI GC buffer high fidelity
1 µl	Oligonucleotide fwrd (1:10 dilution)
1 µl	Oligonucleotide rev (1:10 dilution)
1 µl	dNTPs
0.5 µl	Template DNA
0.25 µl	KapaHIFI-polymerase
36.25 µl	H_2O

Table 16: PCR program used for site-directed mutagenesis

Temperature	Time	Cycles
95°C	5 min	1x
98°C	20 sec	34x
54°C	15 sec	
72°C	30 sec/kb	
72°C	10 min	1x
4°C	∞	1x

Table 17: Reaction-mix for DpnI digest

Volume	Component
50 µl	PCR product
40 µl	H_2O
10 µl	10x NEB4 buffer
2 µl	DpnI enzyme

3.4 Isolation of TraD

3.4.1 Overproduction of TraD

Overproduction in LB medium

Plasmids were transformed as described in Chapter 3.3.4. The transformed cells were inoculated in liquid LB medium with the corresponding antibiotic. The pre-cultures were incubated at 37°C over night in a shaking incubator. 2 L LB medium was inoculated 1:100 with the overnight pre-cultures in 5 L Erlenmeyer flasks. The main-cultures were incubated at 37°C in a shaking incubator at 140 rpm. The growth was monitored by measuring the optical density of the cultures at 600 nm (OD_{600}). At an OD_{600} of 0.6-0.8 the cells were induced with 0.5 mM IPTG and further incubated for 2.5 h with 140 rpm. The cells were harvested by centrifugation (4,500 rpm; 20 min; 4°C; Sorvall F10S rotor; Sorvall RC6+ centrifuge) and the cell-pellets were washed with 10 mM Tris/HCl pH 8.0. After a second centrifugation (same as first centrifugation) the cell-pellets were frozen in liquid nitrogen and stored at -80°C until further use.

Overproduction in autoinduction medium

The transformation and the pre-culture were performed as described in Chapter 3.3.4 and in the previous passage. The overnight pre-culture was inoculated with fresh LB medium with antibiotics and was incubated in a shaking incubator at 30°C or 37°C for 10 h. The main-cultures (1 or 2 L autoinduction medium in 5 L Erlenmeyer flaks) were inoculated 1:100 with the second pre-culture and were incubated while shaking (140 rpm) at 30°C or 37°C for different time periods between 4 h and 16 h. The OD_{600} was measured at different time points to monitor the growth of the cultures.

The cells were harvested by centrifugation (6,000 rpm; 4°C; 20 min; Sorvall F10S rotor; Sorvall RC6+ centrifuge) and the pellets were washed with 10 mM Tris/HCl pH 8.0. The cells were frozen as pellets in liquid nitrogen and were stored at -80°C.

Overproduction with Rifampicin

The transformation and the overnight pre-culture were performed as described previously. The 2 L main-culture (LB medium with the corresponding antibiotic) was inoculated 1:100 with the pre-culture in a 5 L Erlenmeyer flask and was incubated in a shaking incubator at 30°C. At an OD_{600} of 0.6 the cells were induced with 0.5 mM IPTG and were incubated for 30 min at 30°C before Rifampicin (200 µg/ml) was added. Since Rifampicin was solved in DMSO the final concentration of DMSO in the culture was 0.5% (v/v). The cultures were further incubated for 1.5 h at 30°C. The cells were harvested with centrifugation (6,000 rpm; 4°C; 20 min; Sorvall F10S rotor; Sorvall RC6+ centrifuge) and the pellets were washed with 50 mM Tris/HCl pH 8.0 and 150 mM NaCl. After a second centrifugation the cells were frozen as pellets and were stored at -80°C.

3.4.2 Isolation of TraD from inclusion bodies

Cell breakage

Frozen cell pellets were resuspended in 10 mM Tris/HCl pH 8.0 (10 ml buffer/ 1 g cells) and 1 mg DNAseI was added together with 1 mM PMSF. The cells were broken with the cell disruptor (Constant Systems) with a pressure of 2.4 kBar. The breakage was repeated twice. The broken cells were centrifuged for 20 min at 4°C (6,000 rpm; Sorvall F10S rotor; Sorvall RC6+ centrifuge). The supernatants were analyzed by 11% SDS-PAGEs and the pellets were used for the further isolation of TraD from inclusion bodies.

Isolation of inclusion bodies

Pellets from the broken cells were resuspended in 10 mM Tris/HCl pH 8 and were centrifuged again (6,000 rpm; Sorvall F10S rotor; Sorvall RC6+ centrifuge). The wash steps of the inclusion body pellets were repeated three times. Finally the pellets were resuspended in 3 ml of 'Denaturing buffer'. To completely solve the resuspended inclusion bodies the samples were incubated at 37°C for 8 min and vortexed vigourosly. The samples were again incubated for 5 min at 37°C and were vortexed shortly. The denatured protein solutions were centrifuged for 30 min at 4°C (85,000 rpm; Beckman&Coulter TLA 110 rotor; Beckman&Coulter Optima Max-XP Ultracentrifuge). The isolated inclusion bodies were analyzed by 11% SDS-PAGEs and were further purified.

Purification of denatured TraD

Denatured TraD protein was purified with Ni-affinity chromatography via gravity flow. Ni-NTA-agarose from Sigma-Aldrich was used as column material and was activated as described in the manual. The column was equilibrated with 5 CV of 'Purification buffer'. The purification was performed at room temperature. The denatured TraD was bound to the column and the purification was performed with the protocol given in Table 18. The different fractions were analyzed with 11% SDS-PAGEs. The samples were precipitated with TCA before loading onto the gel.

Table 18: Buffer conditions used for the purification of denatured TraD

Buffer	$C_{Imidazole}$	CV	Step
Purification buffer	5 mM	2	wash
Purification buffer	10 mM	2	wash
Purification buffer	20 mM	2	wash
Purification buffer	30 mM	2	wash
Purification buffer	40 mM	2	wash
Purification buffer	50 mM	2	wash
Purification buffer	60 mM	2	wash
Purification buffer	70 mM	2	wash
Purification buffer	80 mM	2	wash
Purification buffer	90 mM	2	wash
Purification buffer	100 mM	2	wash
Purification buffer	200 mM	½	elution
Purification buffer	400 mM	1 ½	elution

3.4.3 Refolding of denatured TraD

A screen for refolding conditions of TraD was performed with purified denatured protein. The conditions which were tested are summarized in Table 19. The protein was mixed in different buffers in a 1:20 ratio. The samples were vigourosly vortexed for a few seconds after mixing and were incubated for at least 1 h or over night at 4°C. After incubation the samples were centrifuged (15 min; 4°C; 13,300 rpm; Heraeus Fresco 17 centrifuge). The supernatant was removed carefully and the pellet was resuspended in 40 µl 1x SDS-PAGE loading dye. 250 µl of the supernatant were precipitated with 10% (v/v) TCA (final concentration) and resuspended in 10 µl 2x SDS-PAGE loading dye. After neutralization the precipitated samples of the supernatant and 10 µl of the corresponding pellet suspension were analyzed by 11% SDS-PAGEs. The samples were incubated at 99°C for 5 min before loading on the gel.

Table 19: Condition screening for the refolding of purified denatured TraD

Buffer	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Tris pH 8.5	-	50 mM	50 mM	50 mM	50 mM	50 mM	50 mM	50 mM	50 mM	-	50 mM	50 mM	50 mM	50 mM	50 mM	50 mM	50 mM
MES pH 6.0	50 mM	-	-	-	-	-	-	-	-	50 mM	-	-	-	-	-	-	-
NaCl	240 mM	240 mM	9.6 mM	9.6 mM	9.6 mM	9.6 mM	9.6 mM	9.6 mM	240 mM	9.6 mM	9.6 mM	9.6 mM	9.6 mM	9.6 mM	9.6 mM	9.6 mM	9.6 mM
KCl	10 mM	10 mM	0.4 mM	0.4 mM	0.4 mM	0.4 mM	0.4 mM	10 mM	0.4 mM	0.4 mM	0.4 mM	0.4 mM	0.4 mM	0.4 mM	0.4 mM	0.4 mM	0.4 mM
$MgCl_2$	2 mM	2 mM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
$CaCl_2$	-	2 mM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Arginine	0.5 M	0.5 M	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
EDTA	-	-	1 mM	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glutathione reduced	1 mM	1 mM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Glutathione oxidized	0.1 mM	0.1 mM	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
DTT	-	-	1 mM	1 mM	1 mM	1 mM	1 mM	1 mM	1 mM	1 mM	1 mM	1 mM	1 mM	1 mM	1 mM	1 mM	1 mM
TritonX-100	0.5 %	0.5 %	0.5 %	0.5 %	-	-	0.1 %	0.5 %	0.5 %	0.5 %	-	-	-	-	-	-	-
DDM	-	-	-	-	0.1 %	-	-	-	-	-	-	-	-	-	-	-	-
Zwittergent 3-14	-	-	-	-	-	-	-	-	-	-	0.1 %	-	-	-	-	-	-
DM	-	-	-	-	-	-	-	-	-	-	-	0.1 %	-	-	-	-	-
UDM	-	-	-	-	-	-	-	-	-	-	-	-	0.1 %	-	-	-	-
Chaps	-	-	-	-	-	-	-	-	-	-	-	-	-	0.7 %	-	-	-
ASB-14	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.5 %	-	-
LDAO	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.1 %	-
Sodium-deoxycholate	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.3 %

3.4.4 Separation of inner and outer cell membranes from *E. coli*

Total cell membrans were isolated from broken cells by ultracentrifugation. Therefore, the broken cells were first centrifuged with low spin to remove unbroken cells and inclusion bodies (6,000 rpm; Sorvall F10S rotor; Sorvall RC6+ centrifuge; 15 min; 4°C) followed by a high spin centrifugation step (1 h; 4°C; 40,000 rpm; Sorvall T647.5 rotor; Beckman&Coulter Optima L-90K Ultracentrifuge). The isolated total membrane pellets were resuspended in 25 mM Tris/HCl pH 8.0 (0.3 ml/OD_{600} 1/$L_{Culture}$) and 0.3 ml resuspended membranes were loaded on a sucrose gradient with 55%, 51%, 45% and 36% steps. The samples and the sucrose solutions were always handled on ice. The sucrose was solved in 25 mM Tris/HCl pH 8.0. The volume of the 55%, 51% and 45% steps was 0.9 ml, whereas the volume of the 36% step was only 0.5 ml. The gradient was centrifuged for 30 min at 4°C (100,000 rpm; Beckman&Coulter TLA 110 rotor; Optima Max-XP Ultracentrifuge). The separated inner and outer membranes were collected and diluted 3-fold with 25 mM Tris/HCl pH 8.0. The diluted membranes were centrifuged for 45 min at 4°C (40,000 rpm; Sorvall T647.5 rotor; Beckman&Coulter Optima L90K

Ultracentrifuge). The pellet was resuspended in 25 mM Tris/HCL pH 8.0 with 10% (v/v) glycerol. The samples were analyzed by 11% SDS-PAGEs and were frozen in liquid nitrogen and stored at -80°C.

3.4.5 Solubilization screen of TraD with isolated cell membranes from *E. coli*

The resuspended membranes were diluted 1:50 with the corresponding buffer and were rotated for 1 h at 4°C. The solubilized samples were afterwards centrifuged for 25 min at 4°C (80,000 rpm; Beckmann&Coulter TLA120.1 rotor; Optima Max-XP Ultracentrifuge). The supernatants were carefully removed and the pellets were resuspended in 25 mM Tris/HCl pH 8.0. The supernatant and the pellet fractions were analyzed by 11% SDS-PAGEs. The tested solubilization conditions are summarized in Table 20 and 21.

Table 20: Buffer condition screening for the solubilization of TraD

Buffer 1		**Buffer 2**		**Buffer 3**		**Buffer 4**	
50 mM	Tris/HCl pH8	50 mM	Tris/HCl pH8	50 mM	Tris/HCl pH8	50 mM	Tris/HCl pH8
10% (v/v)	Glycerol	10% (v/v)	Glycerol	10% (v/v)	Glycerol	10% (v/v)	Glycerol
150 mM	NaCl	150 mM	NaCl	150 mM	NaCl	150 mM	NaCl
1 mM	DTT	1 mM	DTT	1 mM	DTT	1 mM	DTT
		+ 1x Complete EDTA free		+ 1 mM EDTA		+ 1 mM PMSF	
+ detergent							

Table 21: Detergent screening used in combination with the buffer condition screening for the solubilization and the membrane detachment of TraD

Concentration	**Detergent/Additives**
1% (v/v)	DM
1% (v/v)	UDM
1% (v/v)	DDM
1% (v/v)	LDAO
1% (v/v)	Triton X-100
1% (v/v)	Zwittergent 3-14
1% (v/v)	Sodium-deoxycholate

Concentration	**Detergent/Additives**
1% (v/v)	Aminosulfobetain-14
1% (v/v)	Chaps
2% (v/v)	Sodium-laurylsarcosine
100 mM	Na_2CO_3
1 M	NaCl

3.5 Isolation of TraI

The overproduction and isolation of TraI was done with support of J. Koch (technical staff).

3.5.1 Overproduction of MBP-TraI fusionprotein

Plasmid pEMH010 was transformed into *E. coli* BL21 Star cells as described in Chapter 3.3.4. Ampicillin was used at a final concentration of 100 µg/ml and 0.5% glucose was added to the liquid

cultures. The transformed cells were plated on LB-Ampicillin-plates and incubated overnight at 37°C. A 50 ml pre-culture was inoculated from colonies from plates and incubated in a shaking incubator at 37°C over night.

2 L LB-medium containing Ampicillin and glucose were inoculated with 20 ml pre-culture in 5 L Erlenmeyer flasks and were incubated in a shaking incubator (140 rpm) at 37°C. The growth was monitored by measuring the OD_{600}. At OD_{600} = 0.5 the overexpression was induced with the addition of IPTG to a final concentration of 0.1 mM. After induction the cells were grown for 3 h before harvesting. Alternatively a similar protocol was used, but the cells were grown at 30°C till an OD_{600} of 0.6, and then overexpression was induced by the addition of IPTG to a final concentration of 0.1 mM IPTG. After induction the cells were further shaking incubated overnigth at 18°C.

The cells were harvested by centrifugation (6,000 rpm; 4°C; 10 min; Sorvall RC6+ centrifuge; Sorvall F10S rotor), the cell-pellets were resuspended in 50 mM Tris/HCl pH 7.5 and centrifuged again (4.200 rpm; 15 min; 4°C; Heraeus Multifuge 1 s-R). The pellets were frozen in liquid nitrogen and stored at -80°C until further use.

3.5.2 Overproduction of native TraI

Plasmid pJK001 was transformed into *E. coli* BL21 Star cells as described in Chapter 3.3.4. The transformed cells were directly inoculated in 50 ml pre-cultures (LB-medium) with 100 µg/ml Ampicillin and 0.5% glucose. The pre-cultures were incubated in a shaking incubator over night at 30°C. 2 L main-cultures (LB-medium + ampicillin + glucose) were inoculated with 20 ml of the over night pre-cultures and were incubated in a shaking incubater at 30°C. At an OD_{600} of 0.6 the cells were induced with 0.5 mM IPTG and further incubated for 2.5 h at 30°C. The cells were harvested by centrifugation (6,000 rpm; 4°C; 10 min; Sorvall RC6+ centrifuge; Sorvall F10S rotor), the cell-pellets were resuspended in 50 mM Tris/HCl pH 7.5 and centrifuged (4.200 rpm; 15 min; 4°C; Heraeus Multifuge 1 s-R). The cell-pellets were frozen in liquid nitrogen and stored at -80°C.

3.5.3 Isolation of MBP-TraI

Cell breakage

Frozen cell-pellets were resuspended in 'His-tag MBP-TraI A' buffer (5 ml buffer/1 g cell-pellet) and protease inhibitor (*Complete* EDTA free; Roche) was added to a concentration recommended by the manufacturer. All steps were performed on ice. The resuspended cells were disrupted with a French Press Cell (SLM Instruments) with a pressure of 1,100 PSI. The breakage was repeated at least 3 times until the suspension changed slightly in color and started to clarify. The broken cells were centrifuged for 5 min at 4°C (8,000 rpm; Sorvall F10S rotor; Sorvall RC6+ centrifuge). The supernatant was centrifuged again (30 min; 4°C; 40,000 rpm; Sorvall T647.5 rotor; Beckman&Coulter Optima L-90K Ultracentrifuge). The cytosol was used for the further isolation of MBP-TraI.

Ni-affinity chromatography

The purification of the MBP-TraI fusionprotein was performed using a 5 ml HiTrap Chelating HP column (GE Healthcare) on an ÄKTA purifier system. The column was loaded with $NiSO_4$ as described in the manual. The column was equilibrated with 5 CV 'His-tag MBP-TraI A' buffer before the cytosol was loaded at a flowrate of 0.5-1 ml/min. The flowthrough was collected separately. The purification program is given in Table 22. Eluting proteins were detected at 280 nm. The protein containing elution fractions were analyzed by 11% SDS-PAGEs and used for further purification.

Table 22: Protocol for the isolation of MBP-TraI fusionprotein with Ni-affinity chromatography

Step	Volume	Flowrate	Program	Buffer
loading	-	1-2 ml/min	external sample pump	isolated cytosol
elution	70 ml	4 ml/min	$gradient_{Inc}$ 100% B; fractionation (4 ml/tube)	His-tag MBP-TraI A; His-tag MBP-TraI B
➢ maximum pressure: 0.3 MPa				

Size exclusion chromatography

The MBP-TraI containing samples were further purified with a SD200 16/60 HiLoad gelfiltration column (GE Healthcare). The column was equilibrated with 1-2 CV filtered and degased 'Gelfiltration MBP-TraI' buffer before the samples were injected. The samples were centrifuged (13,300 rpm; 15 min; 4°C; Heraeus Fresco 17 centrifuge) prior the injection and were either injected unconcentrated or after concentration with Amicon filtration-units (70,000 Da cut-off; Millipore) as described in the manual. The column was loaded with 2 ml sample-volume and was run at a flowrate of 1 ml/min. Fractions with a volume of 2 ml were collected after an elution volume of 40 ml. Eluting proteins were detected at 280 nm. The protein containing fractions were analyzed by 11% SDS-PAGEs. For storage the proteins were mixed with glycerol to a final concentration of 10% (v/v) before freezing in liquid nitrogen and were stored at -80°C.

TEV cleavage

The MBP-TraI fusionprotein containing fractions from the size exclusion chromatography were incubated with TEV protease to yield native non-fused TraI. MBP-TraI was mixed with TEV protease in gelfiltration buffer with ratios of 100:1, 50:1, 10:1, and 5:1. The samples were incubated at 4°C, at room temperature and at 30°C for 16 h. The same experiment was performed in the presence of 0.5 mM EDTA and 1 mM DTT. The digested samples were analyzed by 11% SDS-PAGEs. Samples incubated at 4°C and at room temperature were pooled together and were tried in a further purification with Ni-affinity chromatography as described previously. The elution fractions from the purification were analyzed by 11% SDS-PAGEs. The cleavage protocol is adopted with changes from Waugh *et al.*(291).

3.5.4 Isolation of native TraI without detergents

Cell breakage

Frozen cell-pellets were resuspended in 50 mM Tris/HCl pH 7.6 (5 ml buffer/ 1 g cells) and protease inhibitor (*Complete* EDTA free; Roche) was added together with a small spatula tip of DNAseI (Roche). The cells were broken with a pre-cooled (4°C) cell disruptor (Constant Systems; IUL) at a pressure of 2.4 kbar. The breakage was repeated 3 times until the suspension turned clearer. The cells were always handled on ice. The broken cells were centrifuged for 20 min at 8,000 rpm at 4°C (Sorvall RC6+ centrifuge; Sorvall F10S rotor). A second centrifugation was performed with the supernatant for 1 h with 40,000 rpm at 4°C (Beckman&Coulter Optima L-90K Ultracentrifuge; Sorvall T647.5 rotor). The isolated cytosol was used for the purification of native TraI.

Anion exchange chromatography

The anion exchange chromatography was performed with a self-made 75 ml Q-Sepharose HP column (GE Healthcare) used on an ÄKTA purifier system. The column was equilibrated with 3-4 CV of 50 mM Tris/HCl pH 7.6. The sample was loaded at a flowrate of 2-4 ml/min and the flowthrough was collected separately. The purification-program is given in Table 23. The second buffer used within the protocol contains 50 mM Tris/HCl pH 7.6 and 1 M NaCl. Eluting proteins were detected with UV light at a wavelength of 280 nm. Protein containing fractions were analyzed by 11% SDS-PAGEs and TraI containing fractions were used for further purification.

Table 23: Protocol for the isolation of native TraI with anion exchange chromatography

Step	Volume	Flowrate	Program	Buffer
loading	-	2-4 ml/min	external pump	isolated cytosol
wash	40 ml	4 ml/min	0% B	50 mM Tris/HCl pH 7.0
elution	360 ml	4 ml/min	$gradient_{Inc}$ 100 B; fractionation (8 ml/tube)	50 mM Tris/HCl pH 7.0; 50 mM Tris/HCl pH 7.6 + 1 M NaCl
➢ maximum pressure: 0.3 MPa				

Hydrophobic interaction chromatography

The further purification of the TraI containing samples from the previous performed anion exchange chromatography was done with a self-made 16 ml Phenyl-Sepharose column (GE Healthcare) used on an ÄKTA purifier system. The column was equilibrated with 5 CV of 'Phenyl-Sepharose TraI A' buffer. The samples were adjusted to a final $(NH_4)_2SO_4$ concentration of 40 mM and were loaded at a flowrate of 1-2 ml/min. The flowthrough was collected separately. The purification-program is given in Table 24. Protein containing elution fractions were precipitated with TCA and analyzed by 11% SDS-PAGEs.

Table 24: Protocol for the isolation of native TraI with hydrophobic interaction chromatography

Step	Volume	Flowrate	Program	Buffer
loading	-	1-2 ml/min	external pump	elution fraction from anion exchange chromatography
wash	16 ml	4 ml/min	40% B	Phenyl-Sepharose TraI A; Phenyl-Sepharose TraI B
elution	160 ml	4 ml/min	$gradient_{Dec}$ 40% B; fractionation (4 ml/tube)	Phenyl-Sepharose TraI A; Phenyl-Sepharose TraI B
➢ maximum pressure: 0.5 MPa				

Hydroxyapatite chromatography

TraI containing elution fractions from the hydrophobic interaction chromatography were adjusted to 30 mM potassium-phosphate buffer pH 7.6 before they were further purified with hydroxyapatite chromatography, which was performed with an 18 ml self-made column (Type I ceramic Hydroxyapatite 20 µm; Biorad) with an ÄKTA purifier system. The column was equilibrated with 5 CV of 30 mM potassium-phosphate-buffer pH 7.6. The samples were loaded at a flowrate of 1-2 ml/min and the flowthrough was collected separately. The protocol of the program is given in Table 25. Protein containing elution fractions were TCA precipitated and analyzed by 11% SDS-PAGEs. The TraI containing samples were further purified by size exclusion chromatography.

Table 25: Protocol for the isolation of native TraI with hydroxyapatite chromatography

Step	Volume	Flowrate	Program	Buffer
loading	-	1-2 ml/min	external pump	elution fractions from hydrophobic interaction chrom.
elution 1	20 ml	4 ml/min	5% B; fractionation (4 ml/tube)	30 mM KPi buffer pH 7.6; 500 mM KPi buffer pH 7.6
elution 2	90 ml	4 ml/min	$gradient_{Inc}$ 20% B	30 mM KPi buffer pH 7.6; 500 mM KPi buffer pH 7.6
elution 3	30 ml	4 ml/min	$gradient_{Inc}$ 100% B	500 mM KPi buffer pH 7.6
➢ maximum pressure: 0.4 MPa				

Size exclusion chromatography

Size exclusion chromatography was performed by loading the samples on an SD 200 16/60 HiLoad gelfiltration column (GE Healthcare) on an ÄKTA purifier system as described in Chapter 3.5.3. The column was equilibrated with 2 CV of 50 mM Tris/HCl pH 7.6 and 150 mM NaCl. TraI containing elution fractions from the hydrophobic interaction chromatography were pooled together and concentrated with Amicons (70,000 Da cut-off; Millipore) as described in the manual. The samples were centrifuged (13,300 rpm; 15 min; 4°C; Heraeus Fresco 17 centrifuge) before injection to the column. Protein containing samples were analyzed by 11% SDS-PAGEs. For storage the proteins were

mixed with glycerol to a final concentration of 10% (v/v) and frozen in liquid nitrogen. The samples were stored at -80°C.

3.5.6 Isolation of native TraI with detergents

Cell breakage

Frozen *E. coli* Bl21 Star cells with overproduced native TraI were resuspended in 5 ml 'Ion exchange TraI detergent A' buffer without DDM per 1 g cell-pellet. Protease inhibitor (*Complete* EDTA free; Roche) was added to the concentration recommended by the manufacturer. The resuspended cells were broken with a French Press Cell (SLM Instruments) as described in Chapter 3.5.3. The broken cells were centrifuged for 1 h at 4°C (15,000 rpm; Sorvall SS34 rotor; Sorvall RC6+ centrifuge). The supernatant was mixed with DDM to a final concentration of 1% (v/v) and solubilized for 2 h on ice before a second centrifugation was performed (40,000 rpm; 4°C; 45 min; Sorvall T647.5 rotor; Beckman&Coulter Optima L-90K Ultracentrifuge). The supernatant was used for further isolation of native TraI.

Anion exchange chromatography

A 5 ml HiTrapQ HP column (GE Healthcare) was equilibrated with 5 CV of 'Ion exchange TraI detergent A' buffer before the supernatant was loaded at a flowrate of 1 ml/min. The flowthrough was collected separately. The protocol of the purification program is given in Table 26. The protein containing elution samples and the flowthrough were analyzed by 11% SDS-PAGEs. The flowthrough was used for further purification steps since TraI could not be detected in the elution fractions and remains completely in the flowthrough.

Table 26: Protocol for the isolation of native TraI in the presence of detergent with anion exchange chromatography

Step	Volume	Flowrate	Program	Buffer
loading	-	1-2 ml/min	external pump	isolated cytosol
elution 1	16 ml	4 ml/min	0%B; fraction (4ml/tube)	ion exchange TraI detergent A
elution 2	100 ml	4 ml/min	increasing gradient till 100%B	ion exchange TraI detergent A; ion exchange TraI detergent B
➢ maximum pressure: 0.3 MPa				

1. Cation exchange chromatography

The TraI containing flowthrough from the anion exchange chromatography was further purified with a cation exchange chromatography with a 5 ml HiTrapSP HP column (GE Healthcare). The column was equilibrated with 5 CV of 'Ion exchange TraI detergent A' buffer before the flowthrough was loaded at a flowrate of 1-2 ml/min. The protocol of the program is equal to the anion exchange

chromatography and is given in Table 26. The protein containing samples were analyzed by 11% SDS-PAGEs.

2. Cation exchange chromatography

The protein containing fractions from the first cation exchange chromatography were pooled together and diluted 3 times with 'Ion exchange TraI detergent A' buffer, before the samples were loaded on the equilibrated column. The protocol for the program is similar to the program of the anion exchange chromatography given in Table 26 with the exception that the first elution gradient is performed in a volume of 4 ml and the second elution step is performed in a volume of 16 ml, to elute a higher concentrated protein. The elution fractions were analyzed by 11% SDS-PAGEs and TraI containing samples were further concentrated (2x) with Amicons (70,000 Da cut-off; Millipore) as described in the manual.

Size exclusion chromatography

The size exclusion chromatography was performed as described in Chapter 3.5.3. The SD200 16/60 HiLoad column (GE Healthcare) was equilibrated with 2 CV 'Gelfiltration TraI detergent' buffer. The sample was centrifuged before the injection to remove possible aggregates. The protein containing elution fractions were analyzed by 11% SDS-PAGEs. The TraI samples were frozen in liquid nitrogen and stored at -80°C.

3.6 Isolation of Yaf

3.6.1 Over-production of His-tagged Yaf and native Yaf

Plasmids encoding the respective *yaf* gene were transformed into *E. coli* BL21 Star cells as described in Chapter 3.3.4. The transformed cells were inoculated in 50 ml pre-cultures (LB-medium; 100 µg/ml Ampicillin; 0.5% glucose) and incubated over night in a shaking incubator (200 rpm/37°C). 2 L LB medium (100 µg/ml Ampicillin; 0.5% glucose) were inoculated in 5 L Erlenmeyer flasks with 20 ml of the over night pre-cultures and incubated in a shaking incubator (140 rpm) at 37°C. The growth of the cells was monitored by measuring the OD_{600}. At an OD_{600} = 0.5-0.8 the overexpression of *yaf* was induced by adding IPTG to a final concentration of 0.5 mM. After induction the cells were further incubated for 2.5 h (37°C; 140 rpm). The cells were harvested by centrifugation (15 min; 6,000 rpm; 4°C; Sorvall RC6+ centrifuge; F10S rotor) and were resuspended in 50 mM Tris/HCl pH 7.5. The resuspended cells were centrifuged again (15 min; 4,200 rpm; 4°C; Heraeus Multiguge 1 s-R) and the cell-pellets were frozen in liquid nitrogen and stored at -80°C.

3.6.2 Overproduction of His-tagged Yaf mutants

The overproduction of Yaf mutants was performed as described in Chapter 3.6.1 with changed growth-temperatures and different induction time points. The pre-cultures were grown at 37°C,

whereas the main-cultures were grown at 30°C until an OD_{600} of 0.5-0.8. The cells were induced at OD_{600} = 0.5-0.8 with 0.1 mM IPTG and were shaking (140 rpm) incubated at 18°C over night. Cell-harvesting and storage was performed as described previously.

3.6.3 Overproduction of Yaf with incorporated selenomethionine

The overproduction of N-terminal His-tagged Yaf with incorporated selenomethionine was performed in *E. coli* B834(DE3) cells transformed with the plasmid pEMH008, as described in Chapter 3.3.4. The cells were grown in M9 minimal medium in the presence of L-selenomethionine (50 $\mu g/L_{culture}$). The cells were incubated in a shaking incubator at 37°C. The growth was monitored by measuring the optical density (OD_{600}). At an OD_{600} of 0.5 the cells were induced with 0.5 mM IPTG and were further incubated for 3 h at 37°C. The cells were harvested with centrifugation (15 min; 4°C; Sorvall RC6+ centrifuge; Sorvall F10S rotor) and resuspended in 50 mM Tris/HCL pH 7.0. The cells were centrifuged again (same as first centrifugation) and the cell-pellets were frozen in liquid nitrogen and stored at -80°C.

3.6.4 Isolation of His-tagged Yaf

Cell breakage

Frozen *E. coli* BL21 Star cells with overproduced His-tagged Yaf were resuspended in 10 ml 'His-tag Yaf A' buffer per 1 g cell-pellet. Protease inhibitor was added as recommended by the manufacturer (*Complete* EDTA free; Roche) together with a small spatula tip of DNAseI (Roche) prior resuspension. The cells were always kept on ice and were broken by sonication (5x 4 min; 2 min break; 50% duty; stage 5-6) until the color of the cells turned slightly darker and the suspension starts to clarify. A low spin centrifugation (Sorvall RC6+ centrifuge; Sorvall F10S rotor; 6000 rpm; 10 min; 4°C) was performed to remove the unbroken cells, followed by an ultracentrifugation step (Beckman&Coulter Optima L-90K Ultracentrifuge; Sorvall T647.5 rotor; 35,000 rpm; 1 h; 4°C). The isolated cytosol was used for the purification of His-tagged Yaf.

Ni-affinity chromatography

His-tagged Yaf was purified by Ni-affinity chromatography using the ÄKTA purifier system (GE Healthcare). A 5 ml HiTrap Chelating HP column (GE Healthcare) was loaded with $NiSO_4$ as described in the manual. The column was equilibrated with 5 CV of 'His-tag Yaf A' buffer prior to protein purification. The isolated cytosol was loaded on the column at a flowrate of 0.5-1 ml/min depending on the amount of cytosol. The flowthrough was collected separately. The purification-program is given in Table 27. The pressure limit was 0.3 MPa. The protein containing elution samples were analyzed by 15% SDS-PAGEs and samples with purified Yaf were pooled together and either directly used for further purification with gelfiltration or frozen in liquid nitrogen and stored at -80°C.

The column was regenerated and stored as described in the manual. For the purification of the different mutants the column was stripped and recharged for each mutant.

Table 27: Protocol for the isolation of His-tagged Yaf with Ni-affinity chromatography

Step	Volume	Flowrate	Program	Buffer
loading	-	1 ml/min	external pump	isolated cytosol
wash 1	10 ml	4 ml/min	5% B; increasing gradient	His-tag Yaf A; His-tag Yaf B
wash 2	10 ml	4 ml/min	10% B; increasing gradient	His-tag Yaf A; His-tag Yaf B
wash 3	10 ml	4 ml/min	12.5% B; increasing gradient	His-tag Yaf A; His-tag Yaf B
elution	15 ml	2 ml/min	gradient 100% B; frac. (2ml/tube)	His-tag Yaf A; His-tag Yaf B

Size exclusion chromatography

The final purification step was performed with a size exclusion chromatography with a SD75 16/60 HiLoad column (GE Healthcare). The column was equilibrated with 2 CV of 'Gelfiltration Yaf (His-tag)' buffer. The Yaf containing samples from the Ni-affinity chromatography were pooled together and were either injected without previous concentration or were first concentrated with Amicons (10,000 Da cut-off; Millipore). Before the injection the samples were filtered through a 0.2 µM PVDF syringe filter. The injection volume was 2 ml and the elution was done with a flowrate of 1 ml/min. After 40 ml fractions of 2 ml were collected. The total elution volume was 1 CV. A pressure limit of 0.3 MPa was used as recommended by the manufacturer. The protein containing samples were analyzed by 15% SDS-PAGEs and were frozen in liquid nitrogen. The samples were stored at -80°C.

3.6.5 Isolation of native Yaf

The isolation of native Yaf was done with support of J. Koch (technical staff).

Cell breakage

Frozen cells with overproduced native Yaf were resuspended in 'Ion exchange Yaf A' buffer with 1x protease inhibitor (*Complete* EDTA free; Roche). The cell breakage and the isolation of the cytosol were performed as described in Chapter 3.6.4. The cytosol was used for the isolation of native Yaf.

Anion exchange chromatography

A 5 ml HiTrapQ HP column (GE Healthcare) was equilibrated with 5 CV of 'Ion exchange Yaf A' buffer before the isolated cytosol was loaded at a flowrate of 1 ml/min. The flowthrough was collected separately. The protocol for the program is given in Table 28. Protein was detected with UV light at 280 nm. The protein containing samples were analyzed by 15% SDS-PAGEs. Samples with Yaf were used for further purification steps.

Table 28: Protocol for the isolation of native Yaf with anion exchange chromatography

Step	Volume	Flowrate	Program	Buffer
loading	-	0.5-2 ml/min	external pump	isolated cytosol
elution	100 ml	4 ml/min	100% B; Fraction (4 ml/tube)	ion exchange Yaf A; ion exchange Yaf B
➢ maximum pressure: 0.3 MPa				

1. Cation exchange chromatography

Yaf containing samples were further purified by cation exchange chromatography. A 5 ml HiTrapSP column (GE Healthcare) was equilibrated with 5 CV of 'Ion exchange Yaf A' buffer. The samples were diluted 3x with 'Ion exchange Yaf A' buffer without NaCl before loading to the column with a flowrate of 1 ml/min. The purification-program is given in Table 29. The protein containing fractions were analyzed by 15% SDS-PAGEs.

Table 29: Protocol for the isolation of native Yaf with cation exchange chromatography

Step	Volume	Flowrate	Program	Buffer
loading	-	0.5 ml/min	external pump	elution from anion exchange
elution	20 ml	4 ml/min	50% B; increasing grad.; Fraction (4 ml/min)	ion exchange Yaf A; ion exchange Yaf B
➢ maximum pressure: 0.3 MPa				

2. Cation exchange chromatography

To concentrate the protein, a second cation exchange chromatography was performed with the same column and buffers as described above. Yaf containing samples were pooled together and diluted 2 times with 'Ion exchange Yaf A' buffer before they were loaded on the equilibrated column. The purification-program was similar to the protocol from the first cation exchange, except that the elution gradient was steeper. The eluted samples were analyzed by 15% SDS-PAGEs and further purified with size exclusion chromatography.

Size exclusion chromatography

Size exclusion chromatography was performed with an SD75 16/60 HiLoad gelfiltration column (GE Healthcare) as described in Chapter 3.6.4 using 'Gelfiltration native Yaf' buffer. 2 ml of the previous concentrated Yaf samples were loaded to the column. Protein containing samples were analyzed by 15% SDS-PAGEs and Yaf containing fractions were frozen in liquid nitrogen and stored at -80°C.

3.7 Analytical methods

3.7.1 Determination of DNA concentration

The concentration of double stranded DNA was determined with a NanoDrop Spectrophotometer (Peqlab) at a wavelength of 260 nm. Contaminations with RNA or proteins were estimated by the ratios of 260/230 nm for RNA and 260/280 nm for proteins. The system was blanked with the corresponding sample buffer prior the concentration determination. The measurements were performed with a sample volume of 1 µl.

3.7.2 Agarose gel electrophoresis

Agarose gel electrophoresis was either performed in SBA buffer for standard DNA analyses or in TB buffer for enzymatic activity assays. Depending on the size of the analyzed DNA, gels containing the corresponding buffer (SBA or TB) with 0.8 – 1% agarose were used. Samples for standard DNA analyses were mixed with 6x Orange DNA loading dye (Fermentas) and run with 140-160 V in SBA buffer. The run time varied with the DNA size, short fragments were run for 20 min whereas plasmids were run for 40 min. The 1 KBplus ladder (Fermentas) was used as size standard. Samples for enzymatic activity assays were treated as described in Chapter 3.8.6. Samples were run with variable time at 100 V in TB buffer in relation to the DNA size, short fragments were run for 30 min and plasmids were run for 1 h.

3.7.3 Visualization of DNA

Double stranded DNA was visualized by intercalation of ethidium bromide and detection of the fluorescence of ethidium bromide with UV-light. To visualize DNA used in cloning procedures the ethidium bromide was directly mixed in a final concentration of 0.01% with the warm agarose-buffer-mix prior casting of the gel. To visualize DNA within enzymatic activity assays the agarose gels were stained afterwards for 15 min in a buffer solution containing 0.1% ethidium bromide. The DNA was visualized and data was documented with a 2UV Transilluminator gel documentation system (UVP).

3.7.4 Determination of protein concentration

The Bradford assay was used to determine the exact protein concentration. The assays were performed with Bradford ready-to-use reagent (Fermentas) as recommended by the manufacturer. Calibration curves were performed with BSA concentrations between 0-2 mg/ml. Each assay included besides a calibration curve a blank sample with the corresponding buffer of the protein. The assays were performed in 96 well-plates with a total reaction volume of 250 µl. The samples were mixed with the Bradford reagent and were incubated for 5 min at room temperature. Afterwards the samples were measured with a plate-reader-spectrophotometer (Infinite M200; Tecan) with a

wavelength of 595 nm. The protein concentrations were calculated by taking the sample volume and the calibration curve into account.

For a fast determination of the protein concentration determination, the protein absorbance of the samples was measured with the NanoDrop Spectrophotometer (Peqlab) at a wavelength of 280 nm. The corresponding buffer was used for a blank sample. The protein concentration was determined based on the extinction coefficient based on the primary amino acid sequence of the protein.

3.7.5 Protein precipitation with TCA

Protein samples were mixed with pre-cooled TCA to a final TCA concentration of 10% (v/v). The samples were shortly vortexted and were incubated on ice for 15 min. After incubation the samples were centrifuged for 15 min at 4°C (13,300 rpm; Heraeus Fresco-17 centrifuge). The supernatants were carefully removed and the pellets were resuspended in 2x SDS-PAGE loading dye. The samples were neutralized with 1 M NaOH until the color of the pH indicator turned from yellow to blue and were then analyzed by SDS-PAGE.

3.7.6 Polyacrylamide gel electrophoresis

SDS-PAGE

SDS-PAGE analyses were performed with the Minigel-system (Biorad) with 0.75 mm gels. The acrylamide content of the separation gels was either 11% for the separation of TraI and TraD or 15% for the separation of Yaf. The compositions of 11% and 15% separation gels are given in Table 30 and 31. The stacking gels were prepared with 4% acrylamide; the composition for stacking gels is given in Table 32. Samples were mixed with 5x protein-loading dye (final concentration 1x) prior loading. TCA precipitated samples were pretreated as described in Chapter 3.7.5. The gels were run in 1x TGS buffer at 30 mA/gel until the blue loading dye front reached the bottom of the gel (~45 min). Gels were stained afterwards with coomassie or silver.

Table 30: Buffer-mix of two 11% separating gels

Volume	Component
2.5 ml	Separation buffer (see Chapter 3.1.3)
3.67 ml	Acrylamide (30%)
3.75 ml	deionized H_2O
60 µl	APS (10% w/v)
7 µl	TEMED

Table 31: Buffer-mix of two 15% separating gels

Volume	Component
2.5 ml	Separation buffer (see Chapter 3.1.3)
5 ml	Acrylamide (30%)
2.38 ml	deionized H_2O
60 µl	APS (10% w/v)
7 µl	TEMED

Table 32: Buffer-mix of two 4% stacking gels

Volume	Component
1.25 ml	Stacking buffer (see Chapter 3.1.3)
0.65 ml	Acrylamide (30%)
3.06 ml	deionized H_2O
60 µl	APS (10% w/v)
7 µl	TEMED

Blue native PAGE

Blue native PAGE analyses were performed with the Minigel-system from Biorad with 0.75 mm gels. Blue native gels were cast as gradient gels with 5-13% acrylamide. 4% stacking gels were layered on top of the native gel. The compositions for 4%, 5% and 13% gels are given in Table 33. The gels were cast at room temperature with a gradient-casting system. The cathode buffer A (high coomassie concentration) was stirred over night at room temperature to achieve complete solubility of the coomassie. The sample buffer was made of cathode buffer A and 10% glycerol (v/v). The size standards used are given in Table 34. The gels were run at 4-7°C. As long as the samples were in the stacking gel the gels were run with cathode buffer A at a constant voltage of 100 V. After entering the native gel the voltage was changed to 150 V. When 1/3 of the gels were reached the cathode buffer was changed to cathode buffer B (low coomassie concentration) and the voltage was changed to 200 V. The molecular masses of the protein samples were calculated based on the migration rates of the marker proteins and the sample. The Blue Native PAGE was performed with support of J. Koch (technical staff).

Table 33: Buffer-mix for the preparation of Blue native PAGE gels

Volume 4% gel	Volume 5% gel	Volume 13% gel	Component
0.6 ml	1.25 ml	3.25 ml	Acrylamide 40%
2 ml	3.33 ml	3.33 ml	Gel buffer (3x)
-	-	2 ml	Glycerol
50 µl	60 µl	60 µl	APS (10% w/v)
5 µl	7 µl	7 µl	TEMED
6 ml	10 ml	10 ml	add deionized H_2O the corresponding volume

Table 34: Marker-proteins used to estimate molecular masses in Blue native PAGEs

Protein	Molecular mass (kDa)
Thyroglobulin	669
Ferritin	440
Aldolase	158
Conalbumin	75
Ovalbumin	40
Ribonuclease	13.7

Denaturing Urea PAGE

Denaturing urea polyacylamide gels were used with the XL gel system of Biorad. The composition for 7.2% gels is given in Table 35. All components, except APS and TEMED, were mixed and heated for 15-30 sec in a microwave to completely solve the urea. Afterwards the solution was filtered through a 0.2 µm sterile filter and was degased until no further air bubbles could be detected. After the addition of APS and TEMED the gels were immediately cast at room temperature. The denaturing urea gels only consist of the separation gel and do not contain a stacking gel.

Prior running samples, the gels were pre-run in 1x TB buffer for 2 h at 300 V. Each pocket was washed with a glass syringe to remove diffusing urea before the samples were loaded. Sample containing gels were run for 1:45 h at 300 V.

Table 35: Buffer-mix for the preparation of denaturing Urea PAGE gels

Volume	Component
17.5 ml	10x TB buffer
31.5 ml	Acrylamide (40%)
73.5 g	Urea
➢ fill up to 175 ml with deionized H_2O	
➢ add prior to casting 52.5 µl TEMED	
➢ add prior to casting 525 µl APS (10% w/v)	

Native TB PAGE

Native TB PAGEs were performed with 0.75 mm Minigels. The composition of 7.5% gels is given in Table 36. After mixing all components the gels were cast at room temperature without stacking gels. Samples were run immediately without a pre-run of the gel. If not otherwise mentioned the gels were run in 1x TB buffer for 1 h at 100 V.

Table 36: Buffer-mix for the preparation of native TB PAGE gels

Volume	Component
9.55 ml	deionized H_2O
1.5 ml	10x TB buffer
3.75 ml	Acrylamide (30%)
105 µl	APS (10% w/v)
18 µl	TEMED

3.7.7 Staining of proteins

Coomassie staining

Polyacrylamide gels were heated in a microwave covered with coomassie staining solution until boiling. The gels were incubated in the hot coomassie staining solution for 5-10 min on a platform shaker and then washed twice in H_2O. The gels were again heated in water and incubated for 5 min with constant shaking. A final destaining step was performed by incubating the gel in hot destaining solution until the background turned colorless. In a final step the gels were washed with H_2O and scanned with an Epson Perfection V700 Photo scanner.

Silver staining

The protocol for silver staining of proteins is adopted from Nesterenko *et al.* (292) and is given in Table 37. All steps were performed at room temperature in petridishes while slowly shaking on a gel shaker (KS-130 basic; IKA). The incubation times in the protocol were used for 0.75 mm polyacrylamide minigels.

Table 37: Protocol for silver staining of proteins in SDS-polyacrylamide gels, adopted from (292).

Step	Solution	Time
fixation	60 ml acetone (stock 50% v/v); 1.5 ml TCA (stock 50% v/v); 25 µl formaldehyde (stock 37% v/v)	5 min
rinse	deionized H_2O	3x5 sec
wash	deionized H_2O	5 min
rinse	deionized H_2O	3x5 sec
pretreat A	60 ml acetone (stock 50% v/v)	5 min
pretreat B	100 µl $Na_2S_2O_3 \cdot 5\ H_2O$ (stock 10% v/v) in 60 ml deionized H_2O	1 min
rinse	deionized H_2O	3x5 sec
impregnation	0.8 ml $AgNO_3$ (stock 20% v/v); 0.6 ml formaldehyde (stock 37% v/v); 60 ml deionized H_2O	8 min
rinse	deionized H_2O	2x5 sec
developer	1.2 g Na_2CO_3; 25 µl formaldehyde (stock 37% v/v); 25 µl $Na_2S_2O_3 \cdot 5\ H_2O$ (stock 10% v/v); 60 ml deionized H_2O	10-30 sec
stop-solution	1% v/v glacial acetic acid in deionized H_2O	30 sec
wash	deionized H_2O	10 sec

3.7.8 MALDI-TOF mass-spectrometry and Nano-LC

The identification of specific protein-bands in SDS polyacrylamide gels was done with MALDI-TOF peptide fingerpint analyses performed by J. Kahnt (department of ecophysiology; Max-Planck-Institute for terrestrial microbiology). The samples were cut from the gels and trypsin digested to generate peptide fragments which were used for the identification with mass-spectrometry. Samples that were difficult to measure with MALDI-TOF mass-spectrometry were additionally analyzed with Nano-liquid-chromatography which was also performed by J. Kahnt.

Samples from limited proteolysis experiments were first extracted with organic solvents to remove NaCl and DDM, which had a disturbing effect on the MALDI-TOF measurement. After extraction the samples were mixed and fixed on the corresponding matrix. The organic extraction and the measurements were performed by J. Kahnt.

3.8 Biochemical methods for protein-characterization

3.8.1 Labeling of oligonucleotides with γ-P^{32}-ATP

The reaction-mix for γ-P^{32}-ATP labeling of oligonucleotides is given in Table 38. The reaction was incubated for 30 min at 37°C and inactivated for 20 min at 65°C. After inactivation the reaction-mix was shortly centrifuged (Heraeus Pico 17 centrifuge) and diluted with 36 µl 10 mM Tris/HCl pH 8.5 to a final concentration of oligonucleotides of 5 µM. The labeled oligonucleotides were stored at -20°C.

Table 38: Reaction-mix for γ-P^{32}-ATP labeling of oligonucleotides

Volume	Component
2 µl	Oligonucleotide
0.4 µl	10x PNK-buffer
0.45 µl	H_2O
0.4 µl	T4-PNK
0.75 µl	γ-P^{32}-ATP

3.8.2 Construction of γ-P^{32}-ATP labeled PCR products

PCRs with γ-P^{32}-ATP labeled oligonucleotides were performed to obtain γ-P^{32}-ATP labeled DNA-fragments for the use in DNA-binding and DNA-cleavage assays. The PCRs were performed as described in Table 39 and 40. The PCR products were purified with the DNA Clean & Concentrator Kit 25 (Zymo) as described in the manual. The DNA fragments were eluted in 40 µl H_2O and were analyzed by 7.5% TB-PAGEs. The DNA was stored at -20°C.

Table 39: PCR reaction-mix to generate γ-P^{32}-ATP labeled PCR products

Volume	Component
10 µl	5x Phusion HF-buffer
2 µl	γ-P^{32}-ATP labeled Oligo 1
2 µl	γ-P^{32}-ATP labeled Oligo 2
1 µl	dNTPs
0.5 µl	MS11 genomic DNA
0.5 µl	Phusion-Polymerase
34 µl	H_2O

Table 40: PCR program used to generate γ-P^{32}-ATP labeled PCR products

Temperature	Time	Cycles
98°C	30 sec	1x
98°C	10 sec	34x
53°C	30 sec	
72°C	15 sec	
72°C	10 min	1x
4°C	∞	1x

3.8.3 DNA binding assay

DNA binding assays were used to analize the DNA binding affinity of purified TraI and Yaf. Different γ-P^{32}-ATP labeled DNA substrates were used in a concentration of 100 nm. The protein concentrations varied within the different assays and are given in the corresponding result part. The labeling of the DNA substrates was performed as described in Chapters 3.8.1 and 3.8.2. The reaction-mix is given in Table 41. The protein and the DNA were added as final components. The total reaction volume was 10 µl. The DNA binding was tested in the presence and the absence of different metal ions. The reaction was incubated at 25°C for 1 h. TB buffer was added to the reaction to a final concentration of 1x after incubation. The samples were shortly spun down and 1.5 µl 5x native loading dye were added before running the samples on 7.5% TB-gels for 45 min with 100 V. The gels were incubated over night with reusable phosphoscreens (Molecular Dynamics). The autoradiographs were scanned with the Storm 840 Phosphoscanner (Molecular Dynamics) at the highest resolution.

Table 41: Reaction-mix for DNA binding assays with γ-P^{32}-ATP labeled DNA substrates

Concentration	Component
100 nM	DNA substrate
variable	Protein
34.5 mM	Tris/HCl pH 7.6
103.5 mM	NaCl

Concentration	Component
6.9% (v/v)	Glycerol
0.03% (v/v)	DDM
5 mM	Metal*

* either Ca^{2+}, Co^{2+}, Cu^{2+}, Ni^{2+}, Mg^{2+}, Mn^{2+}, or Zn^{2+}

3.8.4 DNA binding and competition assay

DNA binding and competition assays were performed with isolated native TraI and different ssDNA and dsDNA substrates. The reaction-mix was incubated with unlabeled T_n nucleotides (T_6 till T_{28} with a difference of 3 nucleotides between each step) for 1 h on ice. The total reaction volume was 10 µl. After 1 h incubation the γ-P^{32} labeled competitor DNA was added and the samples were further incubated for 10 min on ice. The samples were shortly spun down and 2 µl of 6x DNA loading dye were added. The samples were analyzed on 7.5% native TB gels. TraI was used in a final concentration of 0.3 µM. The unlabeled T_n oligonucleotides were used in a final concentration of 10 µM and were 100x higher concentrated than the γ-P^{32} labeled competitor DNA (final concentration 100 nM). The assays were performed with two dsDNA substrates (555 bps PCR product of GGI *oriT*; 425 bps PCR product of *difA* site) and three ssDNA substrates (T_{75}; T_{25}; 75 bps oligomer of GGI *oriT* fwrd). The gels were incubated over night with reusable phosphoscreens (Molecular Dynamics). The autoradiographs were scanned with the Storm 840 Phosphoscanner (Molecular Dynamics) with the highest resolution.

3.8.5 DNA degradation assay

DNA degradation assays were performed with purified dimeric and tetrameric Yaf with different γ-P^{32}-ATP labeled DNA substrates. The total reaction volume was 10 µl and the reaction-mix is given in Table 42. The DNA and the protein were added as final components. The assays were performed with different pH conditions and metal ions. Negative controls were done with inactivated proteins which were incubated 10 min at 95°C prior the degradation assays. The mixed samples were shortly spun down and were incubated for 1.5 h at 37°C. After incubation the samples were again shortly spun down and 1 µl of 5x native loading dye were added. The samples were run on 7.5% TB gels for 45 min with 100 V. The gels were incubated over night with reusable phosphoscreens (Molecular Dynamics). The autoradiographs were scanned with the Storm 840 Phosphoscanner (Molecular Dynamics) with the highest resolution.

Table 42: Reaction-mix for DNA degradation assays with γ-P^{32}-ATP labeled DNA substrates

Concentration	Component
100 nM	ssDNA/ds DNA PCR product
2 µM	Protein
50 mM	Buffer
154.5 mM	NaCl
8% (v/v)	Glycerol
5 mM	Metal

3.8.6 DNA relaxation assay

DNA relaxation assays were performed with purified TraI or Yaf using circular supercoiled plasmid DNA substrates, containing either the GGI *oriT* region or the F-plasmid *oriT* region (pEP073 and pEP074). The total reaction volume was 15 µl. An overview on the reaction conditions is given in Table 43. The reaction conditions changed slightly in assays used for the comparison of the activity of TraI and Yaf to the activity of Topoisomerase I from *E. coli*. The reaction-mix used in these assays is given in Table 44. The protein concentrations used in different assays varied, but the total volume of protein containing samples was limited to 7.5 µl. The protein and the DNA were added as final components to the reaction and the samples were shortly vortexed and spun down. The relaxation reactions were performed for 1 h at 37°C and were stopped by the addition of 0.8 µl EDTA (0.5 M; pH 8.0) and 0.8 µl SDS (10% w/v). The reactions containing EDTA and SDS were incubated at 65°C for 20 min and 3 µl of 5x colorless loading dye were added. The samples were shortly spun down before loading to 1% TB agarose gels. The gels were run for 45 min at 110 V in 1x TB buffer and the DNA was stained and visualized with ethidium bromide as described in Chapter 3.7.3.

Table 43: Reaction-mix for DNA relaxation assays

Concentration	Component
variable	Protein
2 nM (80 ng)	DNA (pEP073 or pEP074)
0.5 x	Gelfiltration buffer (Yaf or TraI)
1 x	TB buffer
5 mM	Metal

Table 44: Reaction-mix for DNA relaxation assays performed with Topoisomerase I

Concentration	Component
variable	Protein
2 nM (80 ng)	DNA (pEP073 or pEP074)
0.5 x	Gelfiltration buffer (Yaf or TraI)
1 x	TB buffer
5 mM	Metal
3 mM	KCl
2 mM	$(NH_4)_2SO_4$
0.01 mM	EDTA
0.06 mM	DTT
+ 3% (v/v)	Glycerol

3.8.7 DNA cleavage assay

To analize the DNA cleavage activity of TraI and Yaf the proteins were incubated with different γ-P^{32}-ATP labeled DNA substrates under variable reaction conditions. The DNA labeling was performed as described in Chapters 3.8.1 and 3.8.2. An overview on the tested conditions is given in Table 45 and the used DNA substrates are given in Table 46. The protein and the DNA were added as final components and the samples were shortly vortexed and spun down. The total reaction volume was 10 µl. The reactions were incubated for 1 h at 25°C. After incubation the reactions were shortly spun down and stopped by the addition of one reaction volume of 2x stop solution. The samples were incubated for 4 min at 99°C and were shortly spun down before 4 µl 6x DNA loading dye were added. The samples were loaded on pre-run denaturing urea gels (see Chapter 3.7.6). The gels were run with 300 V for 1.75 h in 1x TB buffer and were incubated over night with reusable phosphoscreens (Molecular Dynamics). The autoradiographs were scanned with the Storm 840 Phosphoscanner (Molecular Dynamics) with the highest resolution.

Table 45: γ-P^{32}-ATP labeled DNA substrates used in DNA cleavage assays

Substrate	Length
GGI oriT F	75 bp
GGI oriT R	75 bp
F-plasmid oriT F	76 bp
F-plasmid oriT R	76 bp
T_{74}	74 bp
PCR product GGI oriT	550 bp

Table 46: Reaction conditions used in DNA cleavage assays

Reaction mix	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Tris/HCl pH 7.6	105 mM	105 mM	105 mM	105 mM	105 mM	105 mM	105 mM	105 mM	105 mM	105 mM	105 mM	105 mM	105 mM	105 mM	105 mM	105 mM
HEPES pH 7	5 mM	5 mM	5 mM	5 mM	5 mM	5 mM	5 mM	5 mM	5 mM	5 mM	5 mM	5 mM	5 mM	5 mM	5 mM	5 mM
NaCl	100 mM	100 mM	100 mM	100 mM	100 mM	100 mM	100 mM	100 mM	100 mM	100 mM	100 mM	100 mM	100 mM	100 mM	100 mM	100 mM
Glycerol	6 %	6 %	6 %	6 %	6 %	6 %	6 %	6 %	6 %	6 %	6 %	6 %	6 %	6 %	6 %	6 %
DDM	0.025 %	0.025 %	0.025 %	0.025 %	0.025 %	0.025 %	0.025 %	0.025 %	0.025 %	0.025 %	0.025 %	0.025 %	0.025 %	0.025 %	0.025 %	0.025 %
DNA	100 nM	100 nM	100 nM	100 nM	100 nM	100 nM	100 nM	100 nM	100 nM	100 nM	100 nM	100 nM	100 nM	100 nM	100 nM	100 nM
TraI	-	500 nM	500 nM	500 nM	-	500 nM	500 nM	500 nM	-	500 nM	500 nM	500 nM	-	500 nM	500 nM	500 nM
Yaf	-	-	-	500 nM	-	-	-	500 nM	-	-	-	500 nM	-	-	-	500 nM
$MgCl_2$	-	-	-	-	5 mM	5 mM	5 mM	5 mM	-	-	-	-	-	-	-	-
$MnCl_2$	-	-	-	-	-	-	-	-	5 mM	5 mM	5 mM	5 mM	-	-	-	-
$CoCl_2$	-	-	-	-	-	-	-	-	-	-	-	-	5 mM	5 mM	5 mM	5 mM
ATP	-	-	3 mM	3 mM	-	-	3 mM	3 mM	-	-	3 mM	3 mM	-	-	3 mM	3 mM

3.8.8 Quantitative image analysis

Quantitative analyses of autoradiography images were performed with the Software ImageJ. To quantify and compare distinct areas of an autoradiography image a reference area was picked to subtract the background influence from all analyzed areas. All areas had the same size and quantitative values were calculated by defining the positive control as 100%.

3.8.9 Dynamic light scattering

Dynamic light scattering analyses coupled with size exclusion chromatography was performed in close collaboration with Dr. S. Smits from the group of Prof. Dr. L. Schmitt from the Heinrich-Heine-University in Düsseldorf with different oligomeric forms of purified His-tagged Yaf.

3.8.10 Limited proteolysis

The limited proteolysis experiments were performed with support of J. Koch (technical staff).

Limited trypsin digest of TraI

Limited trypsin proteolysis was performed with purified TraI in reaction conditions given in Table 47. The protein was mixed with all reaction components except trypsin and the samples were incubated for 1 h at 37°C, at room temperature and on ice. Trypsin was added afterwards to final concentrations given in Table 49 and the samples were incubated again for 20 min on ice. The total reaction volume was 10 µl. The reactions were stopped by adding 5x SDS-loading dye and 10 min incubation at 99°C. The samples were analyzed with 13% and 15% SDS-PAGEs as described in Chapter 3.7.6.

Table 47: Reaction conditions used within limited trypsin proteolysis expermiments with TraI

Reaction mix	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**
Protein	0.59 µM	0.59 µM	0.59 µM	0.59 µM	0.59 µM	0.59 µM	0.59 µM	0.59 µM	0.59 µM	0.59 µM	0.59 µM	0.59 µM	0.59 µM	0.59 µM	0.59 µM	0.59 µM	0.59 µM
$MgCl_2$	10 mM	10 mM	10 mM	10 mM	10 mM	10 mM	10 mM	10 mM	10 mM	10 mM	10 mM	10 mM	10 mM	10 mM	10 mM	10 mM	10 mM
T_n Primer	-	-	-	-	-	2.5 µM	2.5 µM	2.5 µM	2.5 µM	-	-	-	-	-	-	-	-
***oriT* rev Primer**	-	-	-	-	-	-	-	-	-	2.5 µM	2.5 µM	2.5 µM	2.5 µM	-	-	-	-
ATP	-	-	-	-	-	-	-	-	-	-	-	-	-	5 mM	5 mM	5 mM	5 mM
Trypsin	0 nM	8 nM	16 nM	50 nM	4 nM	8 nM	16 nM	50 nM	4 nM	8 nM	16 nM	50 nM	4 nM	8 nM	16 nM	50 nM	4 nM

Limited ProteinaseK digest of TraI

The limited ProteinaseK digest of TraI was performed with 1.2 µM native TraI and increasing ProteinaseK concentrations. The total reaction volume was 10 µl. The samples were incubated for 30 min at 37°C, followed by an inactivation of the enzymes at 99°C for 10 min. Afterwards the samples were mixed with 5x SDS loading dye and were analyzed on 13-15% gradient SDS gels as described in Chapter 3.7.6.

Limited trypsin digest of Yaf

Purified Yaf (3.4 µM dimer and 4 µM tetramer) was incubated with trypsin (0 µM; 12 µM; 16 µM; 20 µM) (Sigma Aldrich) in total reaction volumes of 10 µl. The volume was filled up with the corresponding gelfiltration buffer. The mixed samples were incubated at 37°C for 30 min and the reactions were stopped by the addition of 5x SDS loading dye and incubation at 99°C for 10 min. The samples were analyzed on 13-18% gradient SDS gels as described in Chapter 3.7.6.

Limited ProteinaseK digest of Yaf

Limited proteolysis with ProteinaseK was performed in the corresponding gelfiltration buffer with purified Yaf (3.4 µM dimer and 4 µM tetramer) and increasing ProteinaseK (Merck) concentrations (0 µM; 0.86 µM; 1.73 µM; 2.6 µM; 3.46 µM; 6.9 µM) in total reaction volumes of 10 µl. The mixed samples were incubated at 37°C for 30 min and the reactions were stopped by adding 5x SDS loading dye and incubation at 99°C for 10 min. The samples were analyzed on 15-20% gradient SDS gels.

3.8.11 Pull-down assay for protein-protein interaction

Protein-protein interactions were tested with pull down assays using His-Mag-Sepharose beads (GE Healthcare). 100 µl bead-solution were 3x washed with 500 µl buffer (40 mM Imidazole; 75 mM NaCl; 25 mM HEPES pH 7.5; 5% Glycerol). The tubes with the bead-solution were put into a magnetic tube-holder and the supernatants were carefully removed. The protein-protein interactions were tested with one protein containing a His-tag and a second protein having no His-tag. 90 µl of a protein mixture containing both putative interaction partners were incubated with the equilibrated magnetic beads. After a short incubation time the supernatants were carefully removed and the beads were 2x washed with 200 µl buffer (40 mM Imidazole; 75 mM NaCl; 25 mM HEPES pH 7.5; 5% Glycerol). The bound proteins were eluted with 100 µl elution buffer (400 mM Imidazole; 75 mM NaCl; 25 mM HEPES pH 7.5; 5% Glycerol). All samples were TCA precipitated and analyzed by SDS-PAGE. The pull down assays were performed in close collaboration with S. Rempel, a bachelor student in our group who was working on TraC.

3.9 Crystallization of Yaf and TraI

The crystallization of His-tagged Yaf, native Yaf and native TraI is performed in close collaboration with Dr. S. Smits from the group of Prof. Dr. L. Schmitt from the Heinrich-Heine-University in Düsseldorf. The protein samples were analyzed with dynamic light scattering coupled size exclusion chromatography before initial crystallization screens were performed. The crystallization of Yaf is done with the dimeric and tetrameric form of the protein.

4. Results

4.1 Computational analysis of the GGI encoded T4SS and related T4SSs

In nature, many different conjugative T4SSs can be found. A phylogenetic analysis of these conjugative T4SSs revealed that they can be divided into different families based on I) targeting factors, exemplified by the different relaxases and coupling proteins, and II) based on their MPF complexes as exemplified by the VirB4 like ATPases. A differentiation of the T4SSs based on the relaxase and coupling protein families resulted in 8 different families (MOB_B, MOB_T, MOB_V, MOB_Q, MOB_P, MOB_H, MOB_F and MOB_C), while also the MPF complexes could be divided into 8 different families (MPF_T, MPF_F, MPF_G, MPF_I, MPF_B, MPF_C, MPF_{FA} and MPF_{FATA}). Remarkably, many different functional combinations of MOB and MPF families have been found (e.g. MPF_T + MOB_P, MPF_G + MOB_F or MPF_G + MOB_H) (111).

Several T4SS related genes encoded with the GGI show homology to the *E. coli* F-plasmid. This homology is especially observed for genes in the *ltgX-ych* region. This region encodes proteins that are mostly involved in the formation and stabilization of the secretion complex. 13 of these genes (TraL, TraE, TraK, TraB, TraV, TraC, TraW, TraU, TrbC, TraN, TraF, TraH and TraG) are ordered in the same orientation as the *E. coli* F-plasmid. The gene of the peptidoglycan transglycosylase LtgX is oriented in a different direction, but represents a functional homologue of Orf169 of the F-plasmid. A phylogenetic analysis of $TraC_{NG}$ shows that these proteins form an MPF complex similar to those that belong to the MPF_F family. Several proteins encoded within in the *ltgX-ych* region are not present in the F-plasmid, e.g. Yag, DsbC, Ybe, Ybi, Ycb, AtlA and Ych. Also the pilin subunit TraA, and TrbI, the protein that circularizes the TraA pilin subunits, are not found in the F-plasmid. These proteins are normally found in T4SSs that contain an MPF complex that belongs to the MPF_T family. A functional analysis has shown that all proteins with homology to proteins encoded on the F-plasmid are essential for ssDNA secretion, but neither TraA and TrbI, nor the Yag, Ybe, Ybi, Ycb, and Ych proteins are essential for ssDNA secretion. Remarkably, although they are not found in the F-plasmid AtlA and DsbC are important for DNA secretion.

The relaxase TraI and the coupling protein TraD are located together in an operon with the small hypothetical proteins Yaf and Yaa (110, 198, 202). The relaxase TraI belongs to the family of MOB_H relaxases that are often associated with members of the MPF_G family (eg. in the ICEHIN1056 integrative conjugative element), but are also found to be associated with the MPF_I, MPF_F, MPF_T, MPF_{FATA} families. The coupling protein TraD clusters in phylogenetic analyses together with other coupling proteins that are always linked to a relaxase of the MOB_H family. These TraD-like coupling proteins show a shared synteny with Yaa, a small hypothetical protein with 4 putative transmembrane domains. The fourth protein in this operon, Yaf shows only little homology to other proteins.

Deletion studies of proteins encoded within the GGI, but outside of the *yaa-ych* region showed that these proteins, except for ParA and ParB did not affect DNA secretion (PhD thesis E. Pachulec, 2010). Both, ParA and ParB show homology to partitioning proteins that normally are involved in DNA partitioning during cell division. VirC1, a ParA homologue encoded on the Ti-plasmid was shown to be necessary for T-DNA transport in *A. tumefaciens*.

Different types of the GGI have been identified in *N. gonorrhoeae* and in *N. meningitidis*. To identify GGI-like T4SSs in other bacteria, all proteins essential for DNA secretion were screened for synteny using the absynte webtool (http://archaea.u-psud.fr/absynte/). Using this approach, several GGI-like T4SSs could be identified in other Proteobacteria. The genetic regions encoding proteins homologous to components of the T4SS of *N. gonorrhoeae* were visually analyzed with the Microbial Genome Viewer 2.0 (http://mgv2.cmbi.ru.nl/genome/index.html).

PSI-Blast (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi) was used to identify homologues of all genes surrounding the GGI-like regions, either for proteins encoded within the GGI or to other proteins. Remarkably, several GGI-like T4SSs were identified in other bacteria. Figure 4 shows the genetic organization of the identified T4SSs in comparison with the T4SSs of *N. gonorrhoeae, N. meningitides* and *Neisseria bacilliformis*. The GGI-like T4SS is not only restricted to *Neisseriales*. Moreover, similar T4SSs were identified in α-, β- and γ-Proteobacteria. Most of the organisms containing a GGI-like T4SS belong to the group of β-Proteobacteria. Furthermore, the two pathogenic γ-Proteobacteria species (*S. enterica* sp. *enterica* serovar Newport str. SL254 and *P. mirabilis* HI4320), as well as the α-Proteobacterium *Novosphingobium aromaticivorans* DSM12444 encode GGI-like T4SSs.

The genetic organization of the *dtr* operon and the *mpf* operons of the identified GGI-like T4SSs is highly conserved. However, the orientation of the *dtr* operon to the *mpf* operons is changed compared to the orientation in *Neisseriales* in most of the identified T4SSs. In *N. gonorrhoeae,* as well as in *N. meningitidis* and *N. bacilliformis,* the *dtr* operon and the *mpf* operons have opposite directions, whereas in almost all of the other GGI-like T4SSs the *dtr* operon and the *mpf* operons are oriented in the same direction.

A comparison of the different MPF regions shows that the TraL, TraE, TraK, TraB, TraV, TraC, TraW, TraU, TrbC, TraN, TraF, TraH and TraG proteins are conserved in most GGI-like T4SSs. Several of the identified GGI-like T4SSs show insertions in the *mpf* operon. For example, the T4SS of *S. enterica* sp. *enterica* serovar Newport str. SL254 has, next to an insertion of a transposase and several other smaller genes between the *dtr* genes and the *mpf* genes, a 6.5 kb insertion between two opposite oriented copies of *dsbC* and an additional insertion of 74 kb between *traN* and *traF*. The insertions within the *mpf* genes of the T4SS of *P. mirabilis*, the second γ-Proteobacterium with a GGI-like T4SS, are located at the same positions like in *S. enetrica*. An insertion of two genes, *ynd/mosA* and *ync/mosB,* is located between *traA* and *dsbC* and most likely does not influence the assembly of the secretion system. A second insertion of 21.5 kb is found between *traN* and *traF*. Since no biochemical data is available for these systems it is unknown whether these systems are functional.

Remarkably, also the *yag*, *dsbC*, *ybe*, *ybi*, and *ycb* genes are present in almost all of the identified GGI-like T4SSs, however, these genes seem to be less conserved than the *tra* genes from the MPF region. Homologous genes of the outer membrane protein Yag are always associated with the gene encoding the lytic transglycosylase LtgX. Homologues of *yag* are not present in all of the identified GGI-like T4SSs. One of the two chromosomal T4SSs and the plasmid (pALIDE201) encoded T4SS of *Alicycliphilus denitrificans* K601, as well as the T4SS of *N. aromaticivorans* plasmid pNL1 do not contain *yag* homologues, but still encode homologues of *ltgX*. Whereas, the T4SSs of *P. mirabilis* and *S. enterica*, as well as the chromosomal T4SS of *N. aromaticivorans* lack in addition to *yag* homologues also *ltgX* homologues. Homologous genes of the disulfide-bond isomerase DsbC are present in all of the identified GGI-like T4SSs, with the exception of *N. bacilliformis*. In almost all of the T4SSs the *dsbC* homologues are associated with *traV*, however, the T4SSs of the γ-Proteobacteria

P. mirabilis and *S. enterica* represent exceptions. Actually, the T4SS of *S. enterica* includes two copies of *dsbC* homologues. The T4SS of *A. denitrificans* plasmid pALIDE201 and the chromosomal T4SS of *Acidovorax Js42* contain homologous genes of *dsbC* also associated with the *dtr* genes, representing an exception to the other T4SSs. Homologues of *ybe* and *ycb* are conserved among almost all hereby identified T4SSs, only the T4SS of *N. bacilliformis* and both T4SSs of *N. aromaticivorans* lack *ybe* homologues, whereas *ycb* homologues are only missing in the T4SSs of *N. aromaticivorans*. Homologues of *ybi* are less conserved among the GGI-like T4SSs and are only present in *N. gonorrhoeae*, *N. meningitidis* and the plasmid encoded T4SSs of *A. denitrificans* (pALIDE201) and *Acidovorax* JS42 (pAOV001). Remarkably, genes encoding for the small putatively secreted protein Exp1 are present in most of the GGI-like T4SSs. The genes with homology to the *exp1* gene are all transcribed in an opposite direction from the *tra* operons and seem to form a border between the T4SS and genes that encode proteins with other functions. An *exp1* knock-out mutant is still able to secrete ssDNA, suggesting that the gene is not involved in the secretion of DNA. Only the T4SSs of *N. aromaticvorans* DSM12444 plasmid pNL1, *P. mirabilis* and *Acidovorax* Js42 lack homologues of this gene. However, since no biochemical data is available for these T4SSs it unknown whether or not these systems are functional. Exceedingly, many of these genes which are encoded together in the operon with the MPF components are conserved between the different species, but seem not to be involved in ssDNA transport in *N. gonorrhoeae.*

The phylogenetic relationship between the different GGI-like T4SSs was further analyzed by comparing NJ trees of representative *mpf* and *dtr* proteins, generated with the neighbor joining algorithm with a bootstrap value of 1000.

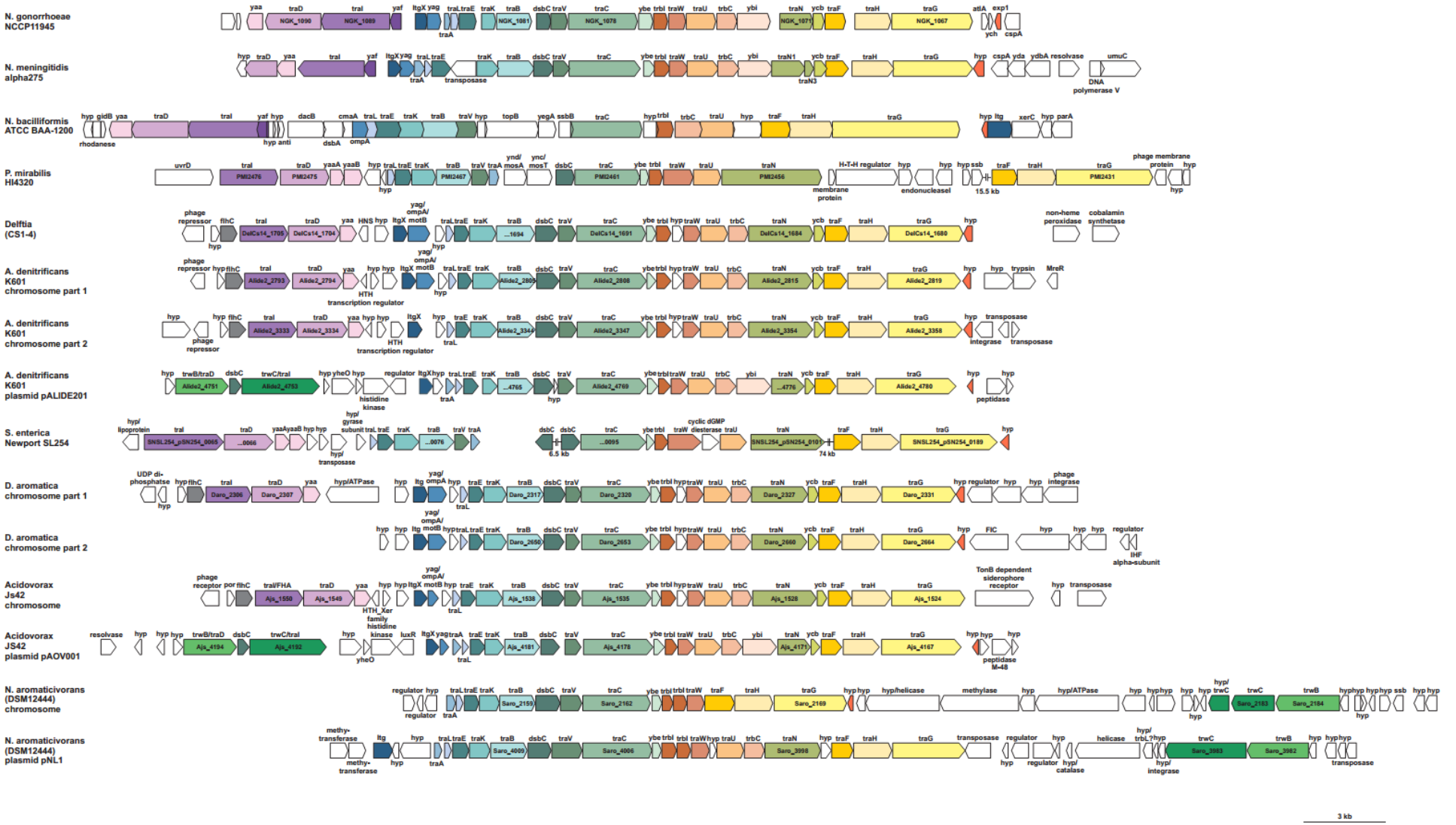
The *dtr* proteins of the GGI-like T4SSs were represented by the relaxases, the coupling proteins, homologues of the small predicted membrane protein Yaa, and homologues of the small predicted protein Yaf. Comparing the phylogenetic trees of all four represented *dtr* proteins shown in Figure 5, the overall clustering of the distinct branches of the trees is similar for TraI, TraD and Yaa. The NJ trees of the analyzed relaxases and coupling proteins essentially show two main branches. The first branch contains of the T4SSs encoded on the plasmids pALIDE201 and PAVO001 of *A. denitrificans* K601 and *Acidovorax* JS42 and the T4SSs encoded on the pNL1 plasmid and the chromosome of *N. aromaticivorans*. The second branch contains all the other organisms. The relaxases and the coupling proteins of the first branch are related to the MOB_F family, while the relaxases and the coupling proteins of the second branch are related to the MOB_H family. As expected, since the *yaa* gene is only found in syntheny with relaxases of the MOB_H family, no homologues of Yaa could be identified for the organisms found in the first branch. This demonstrates that relaxases of two different MOB families can be found associated with a GGI-like MPF complex. Remarkably, all the GGI-like T4SSs encoded on plasmids contained relaxases and coupling proteins of the MOB_F family. This suggests that relaxases of the MOB_H family are preferentially found in T4SSs encoded on the chromosome.

To compare the phylogenetic relationship of the targeting components with the components of the MPF complex, neighbor joining trees were generated for the pilin subunit TraA, the ATPase TraC, the pore forming protein TraB, and the MPF-stabilizing protein TraG. The NJ trees of the four proteins are shown in Figure 6. The four proteins show again essentially very similar trees, suggesting that the proteins in the *mpf* regions from the different organisms have followed the same evolution.

Also the T4SSs encoded on the plasmids pALIDE201 and PAVO001 of *A. denitrificans* K601 and *Acidovorax* JS42 are located close to each other, suggesting that these plasmids have recently been transferred. The T4SSs encoded on the chromosome and on the pNL1 plasmid of *N. aromaticivorans* are again located together, but now they are also located together with the two chromosomal T4SSs of *A. denitrificans* and the T4SSs encoded on the chromosomes of *Delftia sp.* and *Acidovorax* sp. JS42.

It is difficult to draw a conclusion about the evolutionary relationship between the identified GGI-like T4SSs and every assumption has to be seen with caution. One possible scheme for the evolution of these T4SSs is that the ancestral T4SS contained an MPF complex and a targeting system of the MPF_F and MOB_F families respectively. Possibly the MPF_F genes were integrated into the chromosome and acquired there a MOB_H like relaxase and coupling protein, resulting in the genetic organization seen in most of the currently identified systems. In a final step, the MPF_F genes might have been transferred to a plasmid which contained a MOB_F like relaxase and coupling protein. These systems can be found in the pALIDE201, PAVO001, and pNL1 plasmids of *A. denitrificans* K601, *Acidovorax* JS42 and *N. aromaticivorans*.

Figure 4: Comparison of the genetic organization of genome and plasmid encoded GGI-like T4SSs identified in α-, β-, and γ-proteobacteria. The official numbering and nomenclature of the genes is indicated within the gene-arrows. The corresponding homologues are indicated on top of the gene-arrows. T4SS genes are indicated by coloring, homologous genes are labeled in equal colors. The length of the gene-arrows corresponds to the length of the annotated genes. A scale bar of 3 kb is given on the right bottom of the scheme.


N. gonorrhoeae NCCP11945
N. meningitidis alpha275
N. bacilliformis ATCC BAA-1200
P. mirabilis HI4320
Delftia (CS1-4)
A. denitrificans K601 chromosome part 1
A. denitrificans K601 chromosome part 2
A. denitrificans K601 plasmid pALIDE201
S. enterica Newport SL254
D. aromatica chromosome part 1
D. aromatica chromosome part 2
Acidovorax Js42 chromosome
Acidovorax JS42 plasmid pAOV001
N. aromaticivorans (DSM12444) chromosome
N. aromaticivorans (DSM12444) plasmid pNL1
3 kb

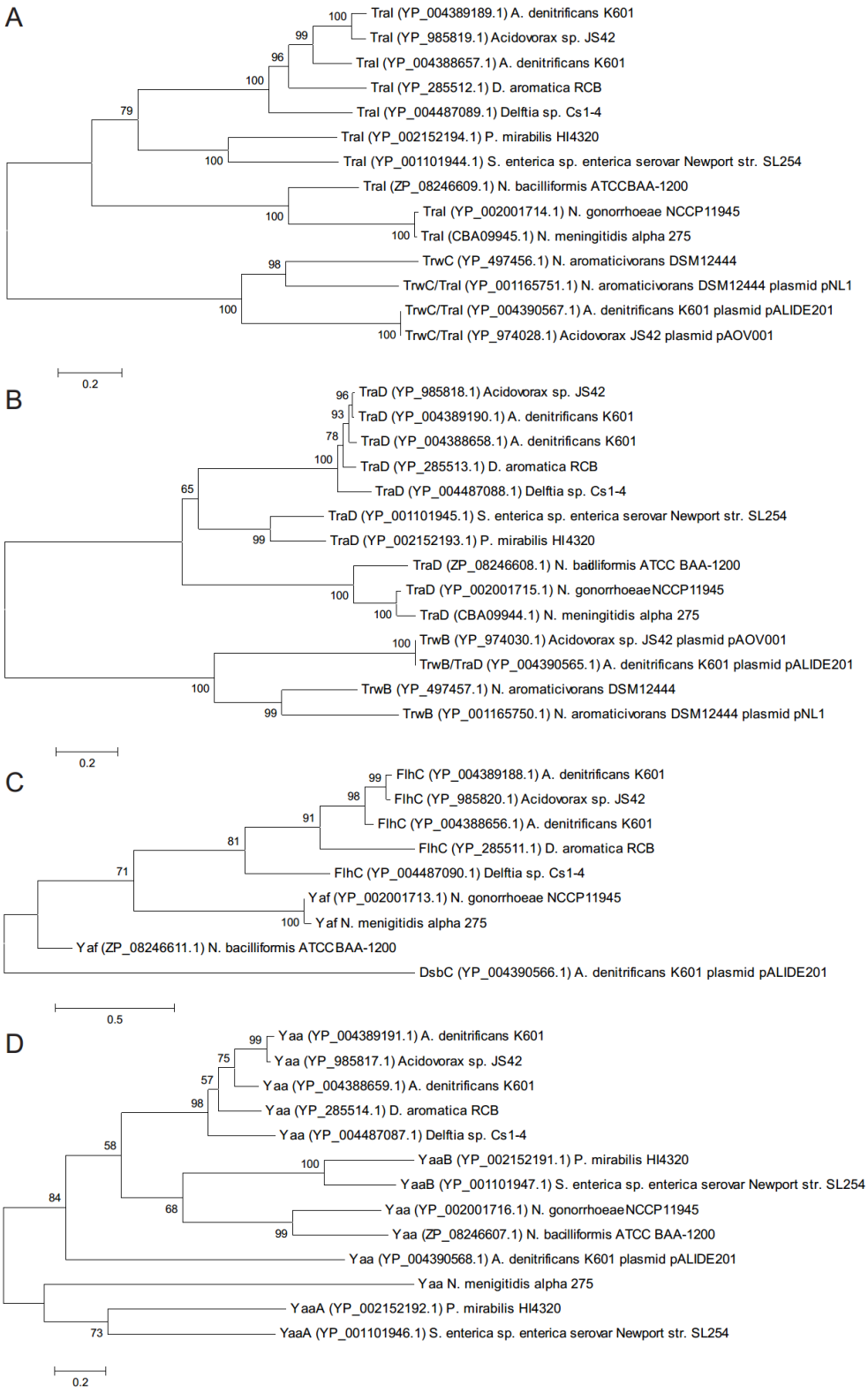


Figure 5: Analyses of the phylogenetic relation between GGI T4SS like homologues of TraI (A), TraD (B), Yaf (C), and Yaa (D). The neighbor-joining algorithm was performed based on amino acid sequence alignments of the homolgous proteins and was generated with a bootstrap value of 1000. The scale bars below the trees indicate the evolutionary distance between the homologues. The numbers at the branching cross-points indicate the percentage of identity of the homologues.

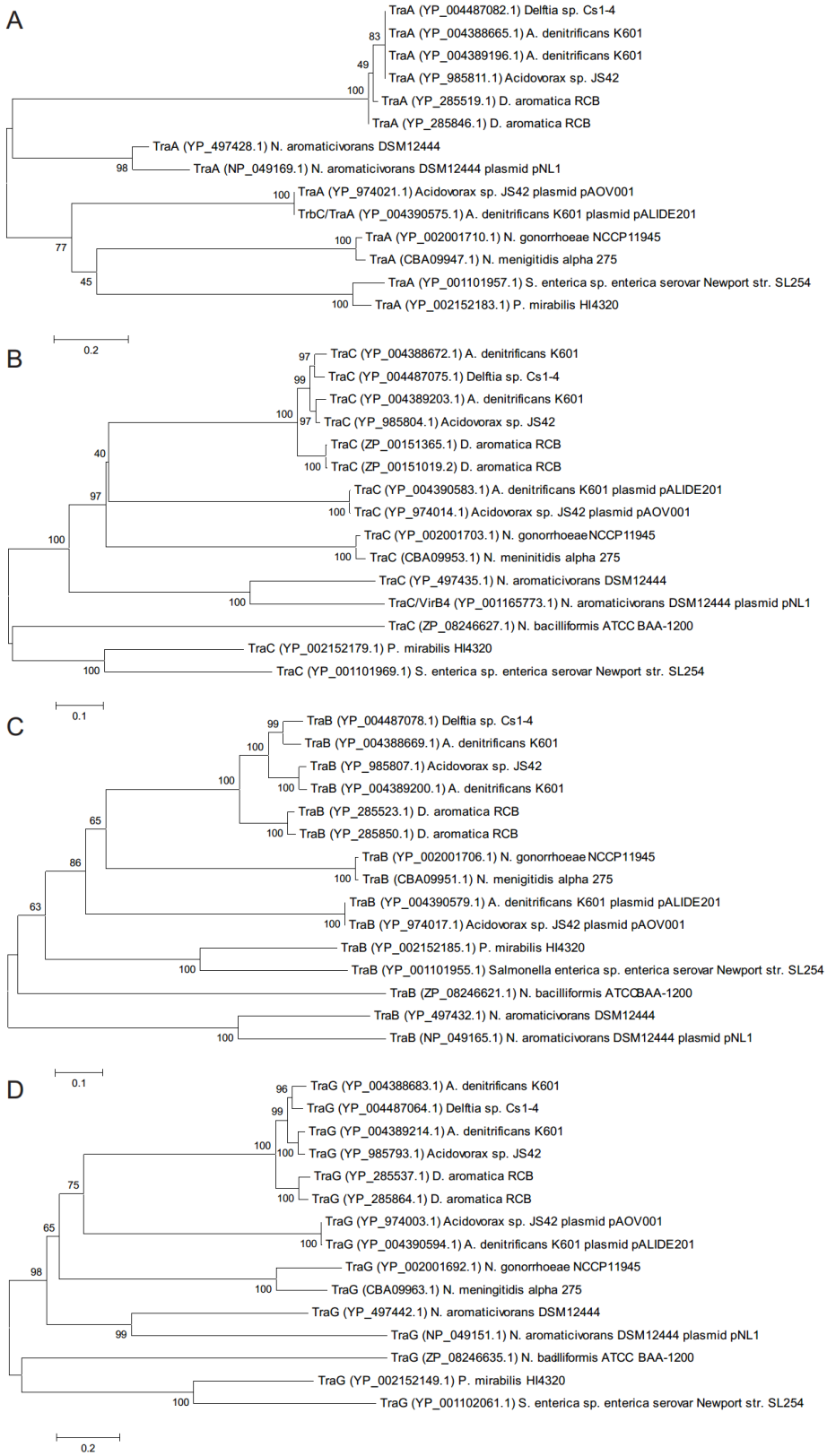


Figure 6: Analyses of the phylogenetic relation between GGI T4SS like homologues of TraA (A), TraC (B), TraB (C), and TraG (D). The neighbor-joining algorithm was performed based on amino acid sequence alignments of the homolgous proteins and was generated with a bootstrap value of 1000. The scale bars below the trees indicate the evolutionary distance between the homologues. The numbers at the branching cross-points indicate the percentage of identity of the homologues.

4.2 Biochemical characterization of the coupling protein TraD

Coupling proteins (CPs) are essential for the function of T4SSs and are present in almost all known T4SSs. They mediate the contact between the secretion substrate and the MPF complex and also play an essential role in targeting the secretion substrates to the secretion complexes. As discussed in Chapter 1.6, the GGI encodes a CP and a relaxase that are related to the MOB_H family of relaxases, but the MPF complex belongs to the MPF_F family. CPs that are related to the MOB_H family, often contain two N-terminal transmembrane domains and are often found associated with a small membrane protein. Like other CPs they contain a nucleotide binding domain (NBD) and an all-α domain (AAD). Computational predictions revealed that the GGI encoded CP also consists of an NBD and an AAD, and that it might also have two N-terminal transmembrane domains. However, both transmembrane domains seem to be less hydrophobic than the transmembrane domains of other members of the VirD4 like CP family. Currently, none of the coupling proteins related to MOB_H like relaxases has been characterized biochemically.

4.2.1 The recombinant overproduction of TraD leads to insoluble protein

To functionally characterize TraD, the protein was overproduced recombinantly in *E. coli* to obtain soluble protein for further purification steps, since the overexpression of proteins in *N. gonorrhoeae* is not well established and further purifications steps are hampered by the fact that they have to be performed under biosafety level 2 conditions.

Initial experiments showed that the overproduction of TraD is a difficult task, since almost all tested overproduction conditions resulted in the production of insoluble and misfolded protein. To optimize the expression level of *traD*, different *traD* expressing constructs either with or without the putative N-terminal transmembrane domains, as well as constructs with different TraD variants fused to the *malE* gene that encodes for the Maltose Binding Protein (MBP) were generated. In the different constructs, *traD* was always cloned behind a T7 promotor. Many attempts with different expression conditions were performed to optimize the overproduction of TraD. Conditions that have been varied included: I) different *E. coli* expression strains like BL21(DE3) and different strains derived from BL21(DE3), like e.g. BL21 Star (for an increased mRNA stability), C43 (for an optimized expression of membrane proteins), BL21(DE3) Tuner and LEMO21(DE3) (for a tunable expression level); II) different growth temperatures (18, 25, 30 and 37°C); III) different media (LB and autoinduction media); III) different IPTG concentrations (0.05 mM - 1 mM); IV) different induction times (1 - 18 h); and V) different temperature shifts (from 37°C to either 30°C, 25°C, or 17°C). Moreover, different conditions were tested in which the *E. coli* strains were co-transformed with a vector that expresses rare tRNAs (pRARE), and it was tried to slow down the production of TraD by adding the RNA-polymerase inhibitor Rifampicin shortly after the induction with IPTG, followed by a cold shock and further incubation of the culture as described in Chapter 3.4.1. In all conditions, the expression of TraD was tested for both, the soluble protein or protein associated with membranes. To distinguish between aggregated proteins in inclusion bodies and membrane proteins, the inner membranes were separated form inclusion bodies by sucrose density step gradient centrifugation as described in Chapter 3.4.4. Table 48 summarizes the results of the optimized expression conditions used with the most promising *traD* constructs. Most conditions resulted in the overproduction of large amounts of

TraD, but unfortunately, non of these conditions resulted in a successful overproduction of large quantities of soluble TraD.

Table 48: Summary of constructs used for the recombinant overproduction of TraD

Protein	Protein amount	Protein soluble	Membrane bound	Inclusion bodies	Purfication successful
TraDΔtm C-terminal His-tag	high	-	partly	high	partly
TraDΔtm N-terminal His-tag	high	-	partly	high	partly
TraDΔtm N-terminal MBP_c-fusion	low	yes	-	low	partly
TraD N-terminal His-tag	high	-	in partly	high	partly
TraD N-terminal His-MBP-TEV fusion	high	yes	-	high	-

The expression of *traD* with constructs that either produce the full-length TraD or a truncated TraD without the transmembrane domain (TraDΔtm) fused to a truncated *malE* gene without signal sequence, resulted in the production of small but detectable amounts of soluble MBP-TraD. Both constructs were generated and overexpressed in close collaboration with T. Deinzer, who performed his bachelor thesis within our group. Even though both constructs resulted in the expression of small amounts of soluble protein, extensive attempts to optimize the purification of these proteins have failed. TraD aggregated on different column materials and could not be eluted from the column after binding. Furthermore, these proteins were inactive in ATP binding and hydrolysis assays. This strongly indicates that these proteins most likely are not folded correctly.

Promising results were obtained with full-length TraD and TraDΔtm grown in autoinduction medium at 30°C. Under this condition still large amounts of inclusion bodies were produced, but TraD could be detected in isolated inner membranes. To optimize the expression conditions for the membrane insertion of TraD, growth experiments with autoinduction medium were performed and inner membranes were isolated at different time points. The large amounts of TraD forming inclusion bodies at the later stages of autoinduction affected the separation of the inner and outer membranes by the sucrose density step gradient centrifugation. Most of the TraD that remains from inclusion bodies formed small pellets at the bottom of the sucrose gradients, but the inner and the outer membranes could not be separated completely from each other. Reloading of the isolated membranes on a second sucrose gradient did not result in either a better separation of inner and outer membranes or a detachment of TraD from the membrane. This suggested that TraD is tightly bound to these membranes and might indicate that full-length TraD is indeed a membrane anchored protein, as it has been observed for other CPs. However, TraDΔtm was detected in similar amounts as full-length TraD in isolated inner membranes derived from TraDΔtm producing strains. Since TraDΔtm lacks the putative transmembrane domains, this result might indicate that TraDΔtm also localizes to the inner membrane in the absence of transmembrane domains. However, it cannot be excluded that the observed proteins in the isolated inner membranes are contaminations derived from inclusion bodies. A comparative SDS-PAGE analysis of isolated inner membranes derived from a

TraDΔtm producing strain collected at different growth times is shown in Figure 7. The overproduced TraDΔtm migrates on SDS-PAGE gels at a position corresponding to a molecular mass of 70 kDa and is dominantly represented in the membrane samples obtained from 12 and 16 h growth time. After further evaluation of the overexpression data (data not shown) the best ratio between the yields of membrane inserted TraD and the begining of the formation of large amounts of inclusion bodies have been determined at a growth time of approximately 10 hours. These conditions were used for further attempts to solubilize and isolate TraD and TraDΔtm from inner membranes of *E. coli*.

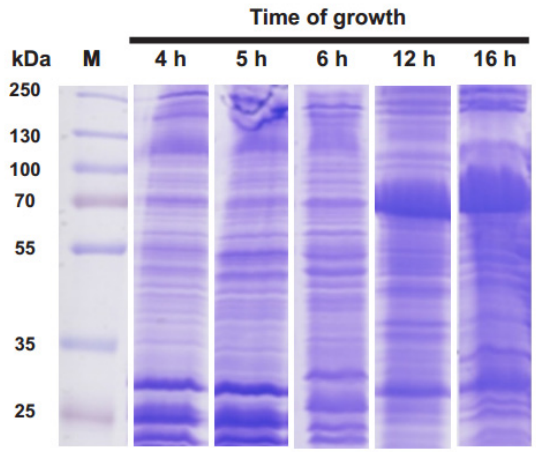


Figure 7: SDS-PAGE with isolated inner membranes of TraDΔtm overproduced in autoinduction medium at 30°C. The corresponding growth time is indicated on top of the lanes. TraDΔtm migrates at a position corresponding to a molecular mass of 70 kDa. A molecular size standard is given on the left site of the gel (PageRuler Plus Protein Ladder; Fermentas).

4.2.2 Membrane-bound TraD cannot be solubilized from isolated inner membranes

The overproduction of full-length TraD and TraDΔtm in autoinduction medium led to an incorporation of the overproduced protein into the inner membranes of *E. coli*. To purify TraD, the overproduced protein should be extracted from the isolated membranes. Therefore, isolated inner membranes were either incubated under conditions with high ionic strength (1 M NaCl or 0.1 M $NaCO_3$; pH 8.0) or in the presence of 10 different detergents. The conditions with high ionic strength have previously been used to extract membrane associated proteins, whereas in the conditions with different detergents the membranes are completely solubilized and also integral membrane proteins are extracted. The 10 different detergents that have been tested to solubilize TraD from the inner membranes included non-ionic, zwitter-ionic and anionic detergents with a range of different critical micellar concentrations. After incubation of the isolated inner membranes, the mixture was centrifuged and the supernatant and pellet fractions from all conditions were analyzed by 11% SDS-PAGEs (see Table 49 and Figure 8). Neither the extraction with 1 M NaCl nor with 0.1 M $NaCO_3$ led to a removal of TraD or TraDΔtm from the membrane. Interestingly, small amounts of solubilized TraD could be observed within 6 different conditions including all tested detergent types (Zwittergent 3-14; DM; LDAO; Laurylsarcosine; ASB-14; Triton X-100). Slight differences were observed comparing the different detergent types with each other. The largest amount of soluble TraD was observed for the two zwitter-ionic detergents Zwittergent 3-14 and ASB-14, as well as the anionic detergent Laurylsarcosine. All three of them are considered as ‘harsh’ detergents. The other three conditions showed less soluble TraD and contained the ‘mild’ non-ionic detergents LDAO, DM and Triton X-100. Even though TraD was solubilzed in part from the inner membranes by using different detergents, further extensive attempts to optimize the purification of TraD have failed. The protein aggregated on different column materials and could not be eluted after

binding. This indicates that the solubilized protein most likely is not folded correctly and suggests that it might represent a contamination of aggregated protein remaining from the inclusion bodies.

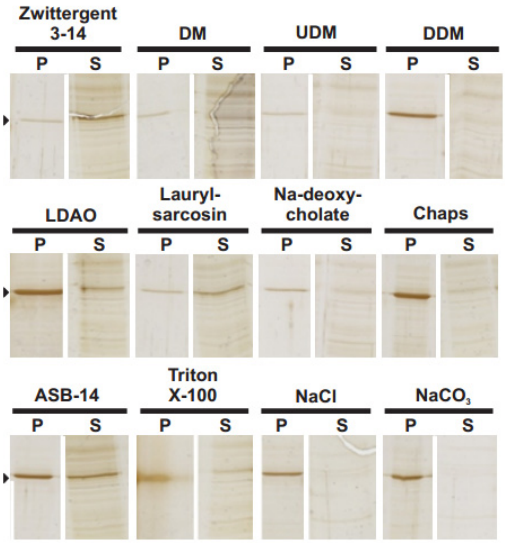


Figure 8: Solubilization and membrane-detachment screen of isolated inner membranes with overproduced TraDΔtm. Quantitative SDS-PAGE analyses with supernatants (S) and resuspended pellet fractions (P) from solubilization and detachment experiments. The corresponding detergents are indicated on top of the gels. The migration position of TraDΔtm is indicated with the black arrows on the left site.

Table 49: Summary of solubilization and membrane-detachment screen with membrane-bound TraDΔtm.
Symbol explanation: + almost all protein; +/- approximately half of the protein amount; - no protein

Detergent	Concentration	Protein insoluble	Protein soluble
Zwittergent 3-14	1 %	+/-	+/-
n-Decyl-β-maltosid (DM)	1 %	+/-	+/-
n-Undecyl-β-D-maltopyranosid (UDM)	1 %	+	-
n-Dodecyl-β-maltosid (DDM)	1 %	+	-
Lauryldimethylamin-N-oxid (LDAO)	1 %	+/-	+/-
Na-laurylsarcosine	2 %	+/-	+/-
Na-deoxycholate	1 %	+	-
Chaps	1 %	+	-
ASB-14	1 %	+	+/-
Triton X-100	1 %	+	+/-
NaCl	1 M	+	-
Na-carbonate	100 mM	+	-

4.2.3 TraD can be isolated and purified as denatured protein

Neither purified TraD nor TraDΔtm could be obtained in an active form after overproduction in *E.coli.* Large amounts of both proteins were, however, available as inclusion bodies. Therefore, it was tried to purify the protein from inclusion bodies. An initial purification of the inclusion bodies was achieved after several wash and centrifugation steps leading to large quantities of insoluble TraD and TraDΔtm. The insoluble protein was denatured in 6 M guanidine and was further purified with

Ni-affinity chromatography as described in Chapter 3.4.2. A representative SDS-PAGE analysis showing the diffent steps of a purification of TraDΔtm is shown in Figure 9. The purified protein migrates at a position corresponding to a molecular mass of 70 kDa and represents the most dominant protein within the samples. The smeary migration pattern of the samples results from remaining guanidine rests which were not removed by TCA precipitation. High losses of protein were observed during the purification due to a weak binding interaction with the column material. This could either result from the choosen buffer conditions or, rather unlikely in denaturing conditions, a suboptimal accessibility of the His-tag. The fact that large amounts of unbound TraD could be detected in the flowthrough also strengthens the assumption of a weak binding affinity. The elution of TraDΔtm started already in the presence of low Imidazole concentrations of 20 mM within the performed wash steps. Several optimization trials to reduce the loss of protein failed. Other elution protocols, like e.g. elution at lower pH also failed and resulted in aggregation of the protein on the column. Despite this, the purification of denatured TraDΔtm in the presence of Triton X-100 resulted in the elution of TraDΔtm of high purity, as shown in Figure 9. Since this condition resulted in reasonable amounts of purified protein, this protocol was used for further isolation of TraD. The purification of denatured protein was succesfull for full-length TraD and TraDΔtm, but the highest amounts and the best purity were obtained with TraDΔtm. Therefore, this protein was used preferably in further refolding assays.

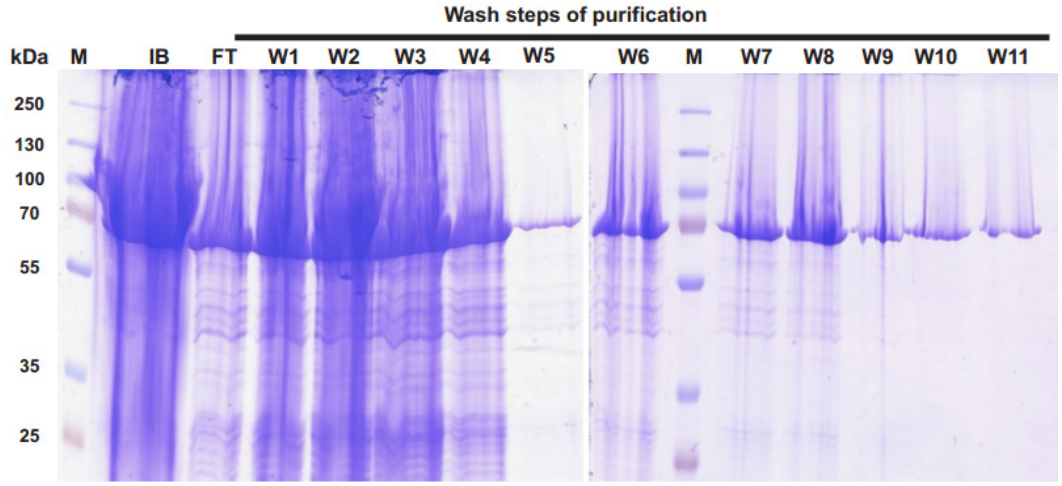


Figure 9: Ni-affinity purification of TraDΔtm under denaturing conditions. 11% SDS-PAGEs with TCA precipitated samples of different purification steps. Molecular masses are indicated on the left site. TraDΔtm migrates at the position of 70 kDa. The labeling of the lanes corresponds the following: M - Prestained Plus Protein Ladder (Fermentas); IB - inclusion bodies; FT - flowthrough; W1 till W11 - different wash steps with increasing Imidazole concentration as described in Chapter 3.4.2.

4.2.4 Soluble TraD can be obtained by refolding the denatured protein

To test whether the purified denatured TraD could be refolded to an active state the purified denatured protein was rapidly diluted into different buffers (buffer conditions are summarized in Chapter 3.4.3 Table 19). After refolding the mixture was centrifuged. The supernatants containing

the soluble and possibly actively refolded protein and pellet fractions of the different refolding conditions were quantitatively analyzed by 11% SDS-PAGEs and are summarized in Figure 10. The results were compared with regards to the solubility of the protein, an immediate pellet formation after mixing the protein with the buffer, and aggregation of the protein in refolding buffer over night. The observed results are summarized in Table 50.

Almost all tested conditions resulted in soluble protein, but differed among each other in terms of the amount of soluble protein. Five conditions showed a slight precipitation of the protein after mixing with the corresponding buffer (‘5’; ‘6’; ‘10’; ‘12’; ‘14’). One refolding condition including Na-deoxycholate (‘17’) immediately ended up in a complete precipitation of the protein after mixing with the refolding buffer. Apperently, the Na-deoxycholate containing buffer and the protein reacted with each other to a white gel-like mixture resulting in an almost complete loss of the protein.

Interestingly, the solubility of TraDΔtm is influenced by the different pH values and detergents. In several conditions TraDΔtm was successfully refolded to soluble protein. The most promising refolding conditions were found with 0.1% (‘7’) and 0.5% (‘8’) Triton X-100, with almost no aggregation of the protein after incubating the samples over night on ice. Condition ‘7’ was used to refold larger volumes of the denatured protein which were further used for purification trials. Unfortunately, the refolded protein could not be purified further since it was not possible to elute the bound protein from the column material. Even testing several different column materials could not solve this problem. Since the refolded protein was already reasonable pure initial ATP binding experiments were performed, but no ATP binding was observed with the soluble protein.

Unfortunately, even after screening many different overexpression conditions, solubilization conditions and refolding conditions, no active protein was obtained. Therefore, a biochemical characterization of the coupling protein TraD was not possible.

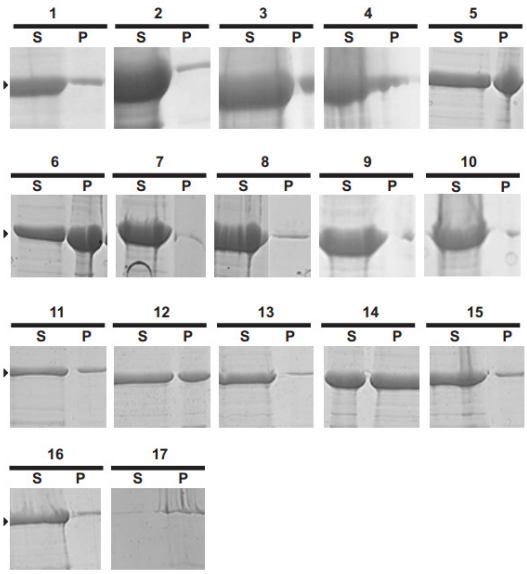


Figure 10: Refolding screen performed with purified TraDΔtm. Quantitative SDS-PAGE analyses of supernatants (S) and resuspended pellet fractions (P) from over night refolding experiments. The buffer conditions are given on top of the lanes corresponding the numbering in Table 50. The migration position of TraDΔtm is indicated with the black arrows on the left site.

Table 50: Summary of refolding screens performed with purified denatured TraDΔtm. The numbering of the buffers correspond the numbering in Table 19, Chapter 3.4.3. Symbol explanation: + almost all protein; +- approximately half of the protein amount; - no protein

Buffer	Protein soluble	Protein precipitated	Protein aggregates
1	+	-	+-
2	+	-	+-
3	+	-	+-
4	+	-	+-
5	+-	+ (white)	+
6	+-	+ (white)	+
7	+	-	-
8	+	-	+-
9	+	-	+-

Buffer	Protein soluble	Protein precipitated	Protein aggragates
10	+	+ (yellow)	+-
11	+	-	+-
12	+	+ (white)	+
13	+	-	+-
14	+-	+ (white)	+
15	+	-	+-
16	+	-	+-
17	-	+ (white)	-

4.3 Biochemical characterization of Yaf

The DNA processing mechanism of T4SSs is often mediated by several different proteins which form together with the DNA a nucleo-protein-complex, the so called relaxosome. The most important component of the relaxosome is the relaxase that introduces a strand specific nick in one strand of the dsDNA. The relaxase is often covalently linked to the nicked DNA strand and is transferred to the recipient cell. In addition to the relaxase, several other proteins are involved in facilitaing DNA processing and targeting, for example by providing access of the relaxase to the DNA, by stabilizing the nucleo-protein complex, by positioning the relaxase in close proximity to the nick site or by facilitating the targeting of the complex to the coupling protein. These proteins are either called nicking accessory proteins or spatial positioning factors.

Many nicking accessory proteins are relatively small proteins (approximately 12-17 kDa) and have a conserved ribbon-helix-helix domain that is often found in transcriptional regulators. To facilitate accessibility of the relaxase to the nick site within the *oriT* region, two homodimers bind to specific DNA sequences near the *oriT* and tetramerize via protein-protein interactions. The formation of a tetramer introduces a bend into the dsDNA that enables the positioning of the relaxase. The overall homology of these proteins is very low and, therefore, these proteins are difficult to identify in genomes or on plasmids using bioinformatical approaches. Possibly, the DNA processing mechanism of the T4SSs of *N. gonorrhoeae* also involves, next to the relaxase TraI, other relaxosome components. However, until now, no such relaxosome components have been identified. Since the hypothetical protein Yaf is encoded within the same operon like *traI* and *traD*, it represents a promising candidate for a putative nicking accessory protein in *N. gonorrhoeae*. Similar to other accessory proteins, Yaf is a small protein and has no homology to other proteins. Furthermore, a possible ribbon-helix-helix domain organization of Yaf was suggested by computational secondary structure predictions. Therefore, the hypothetical protein Yaf was isolated and biochemically characterized to study its possible role in the DNA processing mechanism of the T4SS of *N. gonorrhoeae*.

4.3.1 Recombinant overproduction of Yaf leads to high yields of soluble protein

His-tagged Yaf and native Yaf were overproduced in *E. coli* BL21(DE3) Star cells. After induction with IPTG the cells were further incubated at 37°C and reached optical densities of 2.5 – 3.5. None of the cultures showed retarded growth rates. The expression levels of His-tagged and native Yaf were high and large amounts of soluble protein were produced. The yields of soluble protein were approximately equal for both constructs. The purification from 1 L liquid culture yielded in approximately 25 mg Yaf of high purity.

4.3.2 His-tagged Yaf can be isolated using a two step protocol

The isolation of His-tagged Yaf was performed using a two step protocol including a Ni-affinity chromatography step and a size exclusion chromatography step as described in Chapter 3.6.4. Yaf elutes from the Ni-affinity column at an imidazole concentration of 100 mM. The elution profile is shown in Figure 11 A. All elution fractions were analyzed by 15% SDS-PAGEs. Yaf migrates on SDS-PAGE gels at a position corresponding to a molecular mass of 14 kDa. The calculated molecular mass of monomeric His-tagged Yaf is 17.6 kDa. The difference between these two masses can be explained by the influence of several different factors during electrophoresis that interfere with the migration of proteins. These are for example: different salt concentrations in the samples, the pI of the protein with regard to the pH of the electrophoresis buffer, the denaturing state of the protein with regard to the SDS concentration, or different substances present in the sample which could influence the migration of the proteins. Taking the performed SDS-PAGE analysis into account, elution fractions corresponding to fractions B7 to C7 (see Figure 11 A) were pooled together and the pooled samples were directly loaded on a gelfiltration column. The elution profile from the size exclusion chromatography showed several peaks that are in part difficult to separate from each other (see Figure 11 B). The gelfiltration column was calibrated with commercially available molecular weight standards and the molecular masses of the different elution peaks from isolated Yaf were calculated using the calibration data as reference. A peak with low intensity (25 mAU) was observed at an elution volume of 45 ml which corresponds to the exclusion volume of the column containing aggregated proteins. The maximum of the most dominant protein peak (480 mAU) was observed at 54 ml corresponding to an approximate molecular mass of 69 kDa, which is in agreement with the calculated molecular mass of a Yaf tetramer (70.3 kDa). Another peak eluted with its maximum at 67.5 ml, together with a left and a right shoulder that eluted at 63 ml and 71 ml. The peak at 67.5 ml corresponds to a molecular mass of 29 kDa and most likely represents the Yaf dimer. The shoulders at 63 ml and 71 ml correspond to masses of 38.5 and 23.5 kDa. These peaks most likely correspond with the calculated molecular masses of monomeric Yaf (17.6 kDa) and trimeric Yaf (52.8 kDa). Deviations between the observed molecular masses and the calculated molecular masses can result from a changed migration behavior of the different oligomeric forms on the gelfiltration column, but might also be a result of a slow dynamic equilibrium between the different oligomeric states. The different elution fractions were analyzed by SDS-PAGE and Yaf was detected in each of the fractions confirming the presence of different oligomeric forms of Yaf. The purified protein was free of any detectable contamination on Coomassie stained SDS-PAGE gels and after silver staining only at high protein concentration a second very weak band with a molecular mass of 30 kDa

(possibly the dimeric form of Yaf) appeared, but mass spectrometry analysis did not detect any contamination of the samples.

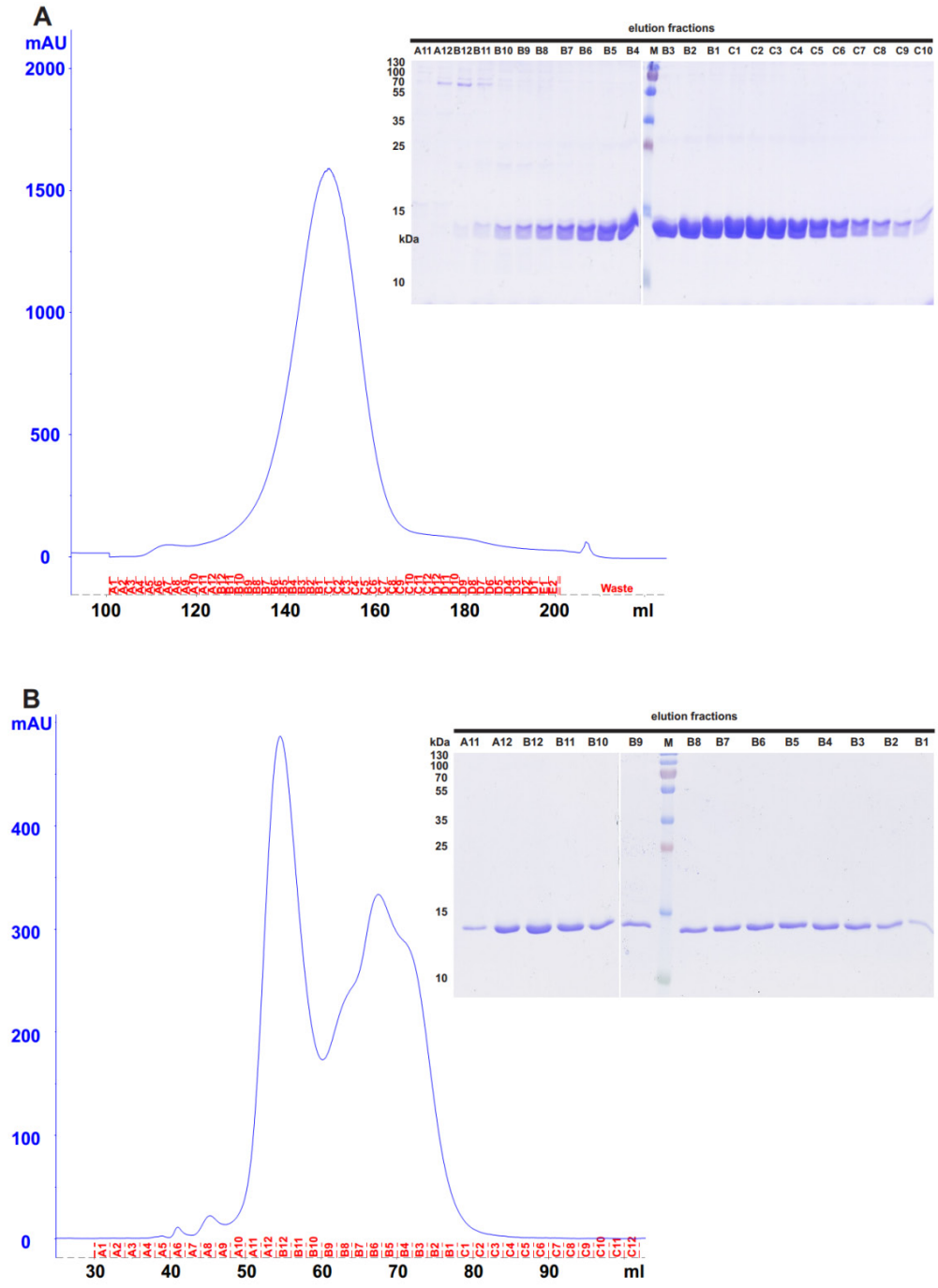


Figure 11: Purification of His-tagged Yaf. (A) Left: Elution profile of Ni-affinity chromatography. The protein was detected at 280 nm and is indicated with the blue line. The intensity is given in mAU. The collected fractions are labeled in red. Right: 15% SDS-PAGEs with elution fractions from the Ni-affinity chromatography. The labeling of the lanes corresponds to the labeling of the elution fractions shown left. Molecular masses are indicated on the left site (PageRuler prestained protein ladder; Fermentas). Yaf migrates at a position corresponding to a molecular mass of 14 kDa. (B) Left: Elution profile from size exclusion chromatography. The labeling and color-code correspond to the description in 'A'. Right: 15% SDS-PAGE with elution fractions from size exclusion chromatography. The lanes represent the corresponding elution fractions. Molecular masses are indicated on the left site (PageRuler prestained protein ladder; Fermentas). Yaf migrates at a molecular mass of 14 kDa.

4.3.3 Native Yaf can be isolated using a four step protocol

The isolation of native Yaf requires a four step protocol including an anion exchange step followed by two cation exchange steps and a size exclusion chromatography step. The protocol is described in detail in Chapter 3.6.5.

Native Yaf eluted from the anion exchange column at NaCl concentrations of 250 – 500 mM (fractions A6 till A11 in Figure 12 A). The elution fractions were analyzed by 15% SDS-PAGEs shown next to the elution profile in Figure 12 A. Yaf eluted in fractions A6 till A11 and migrated at a molecular mass of 13 kDa, which is in reasonable agreement with the molecular mass of a monomer (16.2 kDa) (see Figure 12 A). Several contaminating proteins still remain in the Yaf containing fractions after this first purification step.

Therefore, the Yaf containing elution fractions were pooled and diluted in the buffer from the anion exchange chromatography without NaCl, to enable the binding of the proteins to a cation exchange column in the second purification step. The Yaf containing elution fractions were pooled and again diluted with buffer without NaCl. Yaf was concentrated with a second cation exchange chromatography with a slightly changed protocol. The elution profile from this second cation exchange chromatography, as well as the SDS-PAGEs with different elution fractions are shown in Figure 12 B. Yaf was observed in the peak corresponding to the fractions A5 and A6 (300 – 500 mM NaCl) (see Figure 12 B). The migration patterns of the eluted proteins from the second cation exchange chromatography on SDS-PAGE gels look similar to the results obtained from the anion exchange chromatography and still most of the previously observed contaminations are present.

The Yaf containing fractions were pooled together and were further concentrated before size exclusion chromatography was performed as a final purification step. Similar to His-tagged Yaf, native Yaf showed almost no aggregation during the purification, but differs in the elution profile of the size exclusion chromatography from the His-tagged protein. Native Yaf eluted from the gelfiltration column with two distinct peaks (see Figure 12 C). The first peak had a low intensity (600 mAU) and eluted at a volume corresponding to a molecular mass of 72 kDa. The calculated mass of native Yaf tetramer (64.8 kDa) differs only slightly from the calculated mass of His-tagged Yaf tetramer (70.4 kDa). The second peak had a much higher intensity (> 5000 mAU) and eluted at a volume corresponding to a molecular mass of 33.3 kDa. This peak most likely represents dimeric Yaf (calculated mass of the native dimer is 32.4 kDa). The elution profiles of the size exclusion chromatography of His-tagged and native Yaf clearly show that the presence of a His-tag influences the oligomeric state of Yaf, by stabilizing the tetrameric form. Native Yaf could thus be purified at high concentrations. The fractions from the gel filtration column containing tetrameric Yaf are still contaminated by small amounts of a protein with a molecular mass of 45 kDa.

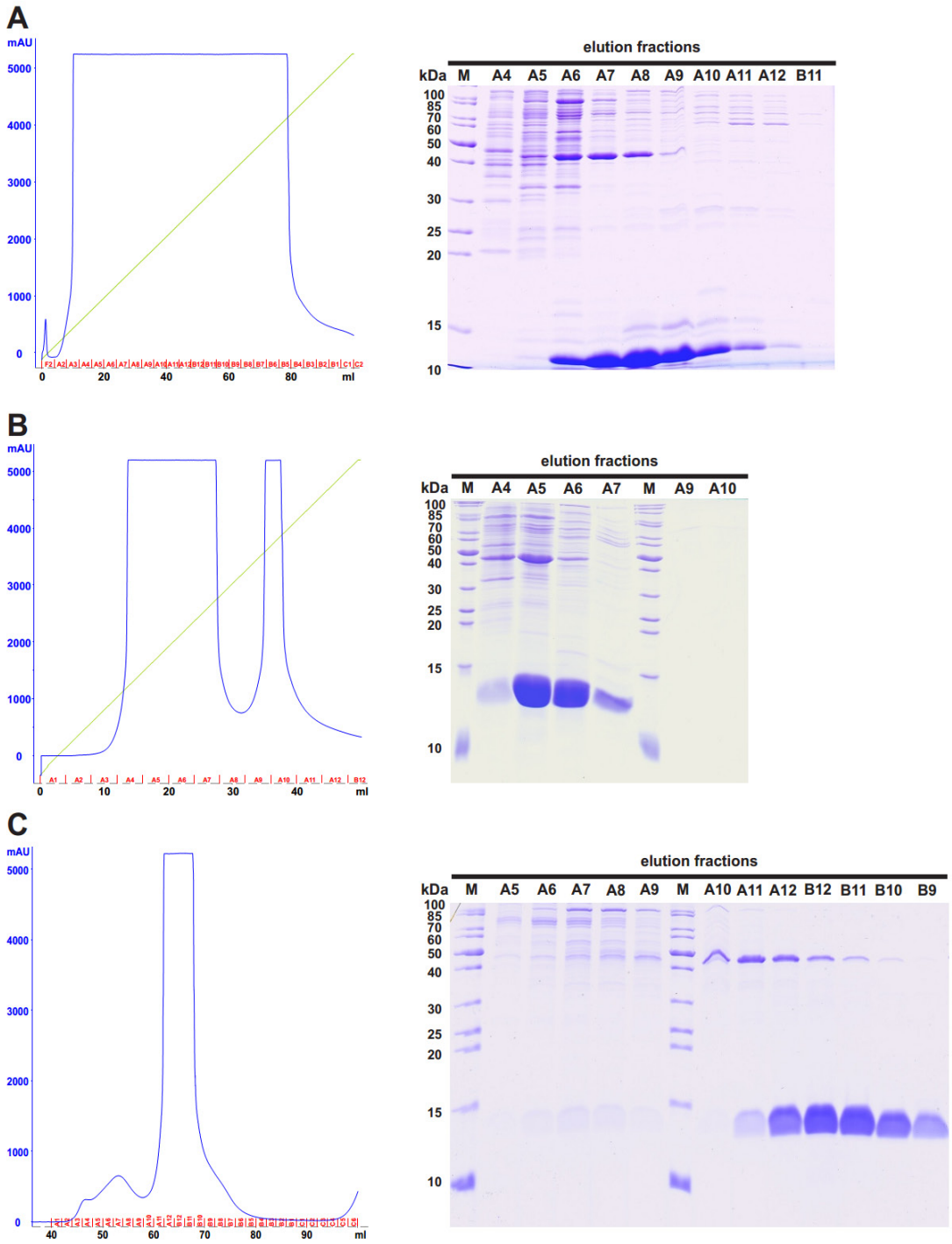


Figure 12: Purification of native Yaf. (A) Left: Elution profile of anion exchange chromatography. The protein was detected with UV light at 280 nm and is indicated as blue line. The intensity of the UV light is given in mAU. The concentration of 'Ion exchange buffer B' is indicated as green line (gradient from 0 – 1 M NaCl). The collected fractions are labeled in red. Middle: 15% SDS-PAGE with elution fractions from anion exchange chromatography. The labeling of the lanes corresponds to the elution fractions. Molecular masses are indicated on the left site (PageRuler unstained protein ladder; Fermentas). Yaf migrates with a molecular mass of 13 kDa within fractions A5 - A10. (B) Left: Elution profile from the second cation exchange chromatography. Right: 15% SDS-PAGE with elution fractions from the second cation exchange chromatography. Yaf migrates with a molecular size of 13 kDa in fractions A5 and A6. The labeling corresponds to the description in 'A'. (C) Left: Elution profile from size exclusion chromatography. Right: 15% SDS-PAGE with elution fractions from size exclusion chromatography. Yaf migrates with a mass of 13 kDa and within fractions A12 - B9. The labeling corresponds to the description in 'A'.

4.3.5 Yaf oligomerizes in four different states with a strong dependency on pH and the presence of a N-terminal His-tag

Several different oligomeric states were observed for isolated His-tagged and native Yaf. The estimated molecular masses of the observed oligomeric forms deviate slightly from the calculated molecular masses and are summarized in Table 51. The most abundant oligomeric forms are dimers and tetramers. The equilibrium between different oligomeric states of Yaf is influenced by the presence of an N-terminal His-tag and by different buffer conditions. The influence of the N-terminal His-tag on the oligomerization of Yaf is depicted in an overlay of the elution profiles from gelfiltrations of His-tagged Yaf and native Yaf (see Figure 13). Both proteins tend to form dimers and tetramers; the His-tagged Yaf mainly forms tetramers, whereas the native Yaf mainly forms dimers, even at high protein concentrations. From these results it can be concluded that the N-terminal His-tag has a stabilizing effect on the formation of tetremeric Yaf, whereas the dimeric form seems to be the naturally oligomeric state of Yaf.

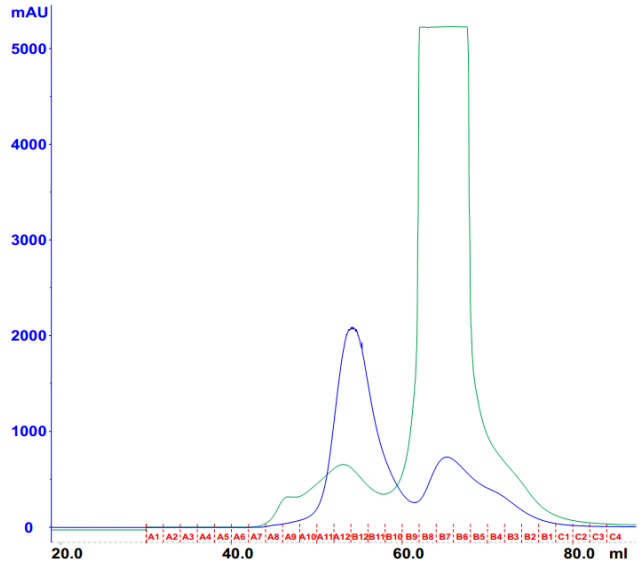


Figure 13: Overlay of elution profiles from gelfiltrations performed with His-tagged Yaf and native Yaf. The elution of the protein was monitored with UV light at 280 nm. The intensities are given in mAU. The collected elution fractions are indicated in red, the elution volume is given in ml below the fractions. The elution profile of His-tagged Yaf is labeled in dark blue, the elution profile of native Yaf is labeled in cyan. The tetramer of Yaf elutes within fractions A11 - B10. The dimer of Yaf elutes within fractions B8 - B4.

Table 51: Overview of the calculated and observed molecular masses of the different oligomeric states of Yaf

Oligomer	$MW_{Calulated}$	$MW_{Gelfiltration}$	$MW_{SDS-PAGE}$
$Monomer_{His-tagged}$	17.6 kDa	20 kDa	14 kDa
$Dimer_{His-tagged}$	35.2 kDa	34.7 kDa	28 kDa
$Trimer_{His-tagged}$	52.8 kDa	48 kDa	-
$Tetramer_{His-tagged}$	70.4 kDa	74 kDa	-
$Monomer_{Native}$	16.2 kDa	-	13 kDa
$Dimer_{Native}$	32.4 kDa	33.3 kDa	-
$Trimer_{Native}$	48.6 kDa	-	-
$Tetramer_{Native}$	64.8 kDa	72 kDa	-

To test whether the oligomeric state of Yaf is further influenced by different buffer conditions, His-tagged Yaf was transferred to different buffers and gelfiltration experiment were performed in the same buffer conditions. The buffer exchanges were performed with a desalting column (GE Healthcare) as described by the manufacturer and the samples were concentrated before injection to the gelfiltration column. To exclude an influence of different buffer components only one component was changed in each experiment. Figure 14 shows an overlay of the elution profiles from different gelfiltration experiments performed under different pH conditions.

Incubation of His-tagged Yaf at pH 6.0 led to significant aggregation of the protein. The non-aggregated protein showed a significant change in the observed elution volumes compared to the protein at pH 7.0. At pH 6.0 His-tagged Yaf eluted at volumes corresponding to molecular masses of 81 kDa and 43 kDa. Most likely these two peaks still correspond to the tetrameric and dimeric forms of Yaf. Possibly the lower pH influences the interaction of the protein with the column material which might lead to a faster elution of the protein from the column then expected.

At pH 7.0, only very little aggregation was observed during the concentration steps and the size exclusion chromatography. Three oligomeric states were identified (see Figure 14; dark grey line). One peak (2100 mAU) eluted at 53.5 ml (corresponding to a mass of 71 kDa) and fits almost perfect with the calculated mass of a tetramer. A second peak (750 mAU) eluted at 65 ml and corresponds with a mass of 34.7 kDa to the size of a dimer (35.2 kDa). A small shoulder (250 mAU) was observed at 71 ml, corresponding to a molecular mass of 23 kDa. The observed mass deviates from the calculated mass but most likely represents the monomer. At pH 7.0 the tetramer is clearly the most abundant oligomeric form.

At pH 8.0 the elution profile shows two distinct peaks with almost no aggregation (see Figure 14; light blue line). Similar to pH 7.0 a peak (2600 mAU) with tetrameric His-tagged Yaf eluted at 53.5 ml. The second observed peak (1450 mAU) eluted at a volume of 73 ml, corresponding a molecular mass of 20 kDa representing the approximate size of a Yaf monomer (17.6 kDa). The oligomerization pattern of Yaf observed at pH 8.0 differed significantly from that observed for pH 6.0 and 7.0, with the loss of the dimeric form and the appearance of the monomeric form.

At pH 9.0 three different oligomeric states can be observed (see Figure 14; violet line). The first elution peak with a volume of 53.5 ml represents tetrameric Yaf, as already observed for the other tested conditions. At a volume of 60 ml a second peak eluted at a volume corresponding to a molecular mass of 48 kDa. The third peak contained monomeric Yaf with a molecular mass of 20 kDa and eluted at a volume of 73 ml. The oligomeric form of a dimer was not observed at pH 9.0. The identity of the peak that eluted at a volume corresponding to a molecular mass of 48 kDa is unclear, but the size suggests that this peak might be related to the trimer. Taken together, these results clearly show that the oligomeric state of Yaf is strongly influenced by the pH.

Additionally, it was analyzed whether the equilibrium between the different oligomers was static or dynamic. To test this, purified Yaf was concentrated using an Amicon concentrator and the oligomerization state was analyzed by size exclusion chromatography. A representative comparison of the elution profiles from gelfiltrations with two different Yaf concentrations is depicted in Figure 15 A. It can be clearly observed that an increasing protein concentration results in a shift from the lower oligomeric states to the higher oligomeric states. This indicates that Yaf is indeed in a dynamic equilibrium between the different oligomeric states. During these experiments it was also

observed that the elution profiles of Yaf differed when they were injected at different times after the concentration. This indicates that the dynamic conversion between the different oligomeric states is a slow process. Fractions corresponding to dimeric and tetrameric Yaf were immediately frozen at -80°C after elution from the gelfiltration column. After thawing the samples were used as quickly as possible to prevent conversion between the different oligomeric states as much as possible.

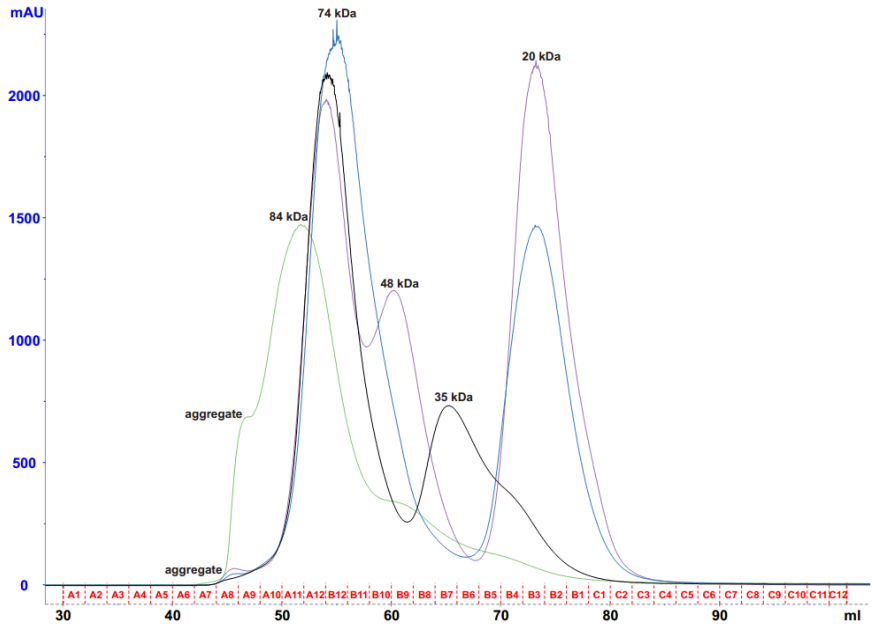


Figure 14: Overlay of elution profiles of gelfiltrations performed in buffers of different pHs. The elution of the His-tagged Yaf was monitored at 280 nm. The elution fractions are indicated in red, the elution volume is given in ml below the fractions. The molecular masses of the eluted proteins are indicated on top of the peaks. Color code of elution profiles from the different tested buffering conditions: green - pH 6; dark grey - pH 7; light blue - pH 8; violet - pH 9

Finally, it was tested whether the presence of the divalent cation Mg^{2+} has an influence on the oligomerization state of Yaf. Therefore, Yaf was transferred to a buffer containing 10 mM $MgCl_2$ and the protein was loaded on a gelfiltration column equilibrated with the same buffer. The overlay of both profils is shown in Figure 15 B. Since no significant effect of $MgCl_2$ was observed, it can be concluded that Mg^{2+} does not influence the oligomerization state of Yaf.

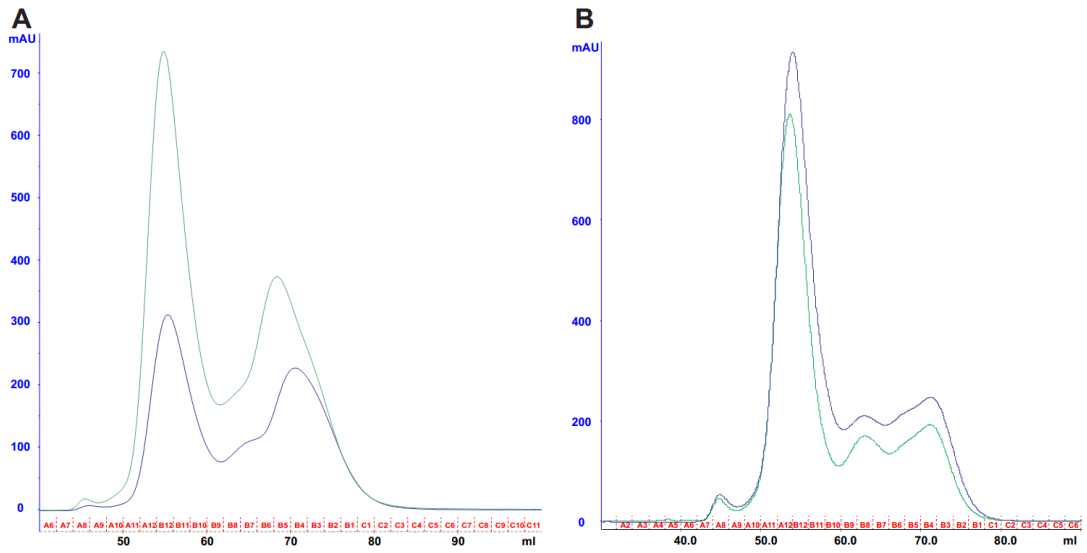


Figure 15: Oligomerization of His-tagged Yaf at different protein concentrations and in the presence of $MgCl_2$. The elution of the proteins was monitored at 280 nm. The collected elution fractions are indicated in red, the elution volume is given in ml below the fractions. (A) Overlay of elution profils with His-tagged Yaf injected at 2 mg/ml (blue) and 4 mg/ml (green). (B) Overlay of elution profiles from His-tagged Yaf without $MgCl_2$ (blue) and in the presence of 10 mM $MgCl_2$ (green).

Interestingly, long term storage of dimeric Yaf at 4°C led to the formation of SDS stable dimers. After storage of dimeric Yaf for > 2 months at 4°C a novel band was reproducibly observed at 28 kDa corresponding to the molecular mass of Yaf dimers. A comparison of the same elution fraction with freshly purified protein and long term stored protein is shown in Figure 16. The samples were mixed with sample buffer containing DTT before loading on the gel. The precise reason why the dimeric state of Yaf is stabilized over time is not known. Most likely, it is caused by the formation of a stabilized form of the dimer over time. This stable dimer was not observed with samples stored frozen at -80°C.

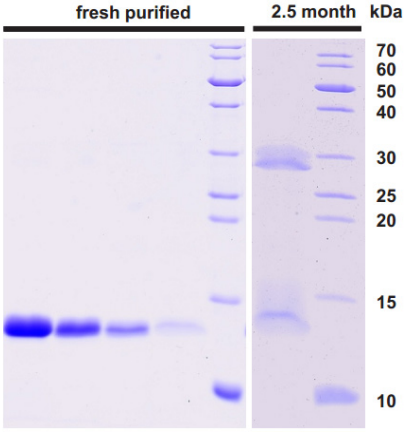


Figure 16: Comparison of freshly purified Yaf dimer and a sample after 2.5 month storage at 4°C. The samples were analyzed with 15% SDS gels. Molecular masses are indicated on the right site (PageRuler unstained protein ladder; Fermentas). Fresh purified Yaf migrates at a position corresponding to a mass of 14 kDa. The 2.5 month old Yaf migrates at positions corresponding to masses of 14 kDa and 29 kDa.

4.3.6 Limited proteolysis revealed a high resistance of Yaf against trypsin and elucidates a possible domain organization of Yaf

To obtain insights into the structural organization of Yaf, limited proteolysis experiments with trypsin and ProteinaseK were performed. The experiments were performed as described in Chapter 3.8.10. The samples were analyzed by SDS-PAGE and the obtained results from both proteolysis experiments are shown in Figure 17. For these experiments elution fractions corresponding to the dimeric and tetrameric forms of Yaf were used.

Remarkably, tetrameric His-tagged Yaf was not digested in the presence of 20 µM trypsin, even after 30 min incubation at 37°C. This demonstrates that the His-tagged tetrameric Yaf is highly resistant to trypsin. The two proteins migrating at molecular masses of 14 and 12 kDa (indicated with '3' on the left site in Figure 17 A) are also present in the trypsin control, indicating that they are most likely either contaminations of the trypsin or autocatalytic cleavage products.

A proteolytic digest of dimeric His-tagged Yaf seems to start at trypsin concentration of 12 µM (see Figure 17 A) suggesting that dimeric His-tagged Yaf is more sensitive to trypsin then tetrameric His-tagged Yaf. Indeed, also dimeric native Yaf was more sensitive to the proteolytic activity of trypsin than tetrameric His-tagged Yaf and was digested at lower trypsin concentrations (data not

shown). This suggests that either the N-terminal His-tag itself prevents the proteolytic digest of Yaf, or that tetrameric Yaf is more resistant to the proteolytic activity of trypsin, possibly due to a restricted accessibility of the protease.

Similar results were obtained for limited proteolysis with ProteinaseK. Like observed with trypsin, dimeric Yaf was also in the presence of ProteinaseK more sensitive to the proteolytic activity, whereas tetrameric Yaf was only digested at higher protease concentrations (see Figure 17 B). Yaf was more sensitive to the proteilytic activity of ProteinaseK than to trypsin, and even the tetrameric Yaf was digested. A digested fragment migrating with a molecular mass of 10 kDa was observed at ProteinaseK concentrations of 0.9-1.7 µM. Such a fragment was not detected in the ProteinaseK control. Thus, the obtained 10 kDa fragment most likely represents a structurally stabilized domain of Yaf.

Based on these results, it can be concluded that Yaf has a structural organization that is higly resistant to digestion with the endopeptidase trypsin, and is more sensitive to the exopeptidase ProteinaseK. Yaf is a very stable protein, and the tetrameric form of Yaf seems to be more stable to proteolytic digestion than the dimeric Yaf. These results also confirm that the dimeric and tetrameric states of Yaf are maintained after rapid freezing of the protein in liquid nitrogen and storage at -80°C, and that these oligomeric states only interconvert slowly.

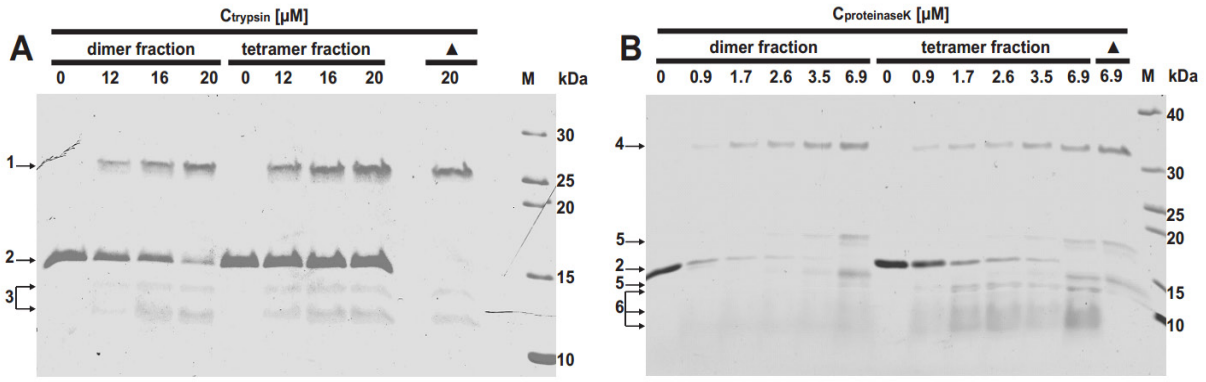


Figure 17: Limited proteolysis of isolated dimeric and tetrameric Yaf. (A) Limited trypsin digest with increasing trypsin concentrations incubated with 3.4 µM dimeric Yaf and 4 µM tetrameric Yaf. The lanes are labeled with the corresponding trypsin concentrations and the digested Yaf oligomer. A typsin control is indicated with the black triangle. Molecular masses are indicated on the right site (PageRuler unstained protein ladder; Fermentas). (B) Limited ProteinaseK digest with increasing ProteinaseK concentrations incubated with 3.4 µM dimeric Yaf and 4 µM tetrameric Yaf. The labeling of the lanes and the molecular size standard correspond to the description in 'A'. Proteins and peptide fragments indicated on the left site: 1-trypsin; 2-Yaf; 3-contamination/self digest products of trypsin; 4-ProteinaseK; 5-contamination/self digest products of ProteinaseK; 6-putative fragments of digested Yaf

4.3.7 Yaf binds to dsDNA with low affinity

The DNA binding activity of Yaf was tested with electrophoretic mobility shift assays with different dsDNA and ssDNA substrates as described in Chapter 3.8.3. Increasing concentrations of Yaf

(0.1-5 µM) were incubated with different γ-P^{32}-ATP labeled PCR products containing either the GGI *oriT* sequence or the GGI flanking *difA* site and were analyzed for DNA binding with native TB PAGEs.

The binding of dsDNA substrates was tested with two different dsDNA substrates. At protein concentrations above 0.5 µM a higher migrating DNA fragment was observed, suggesting that Yaf binds dsDNA with low affinity (> 5 µM). The higher migrating DNA fragment was observed for both dsDNA substrates suggesting that the binding is sequence unspecific (see Figure 18 A). The dsDNA binding was observed in elution fractions corresponding to dimeric and tetrameric forms of Yaf.

Similar DNA binding assays were performed with ssDNA substrates. Even though several ssDNA substrates were tested with different Yaf concentrations, no ssDNA binding was observed. The results from DNA binding experiments with elution fractions containing dimeric and tetrameric Yaf and two different ssDNA substrates are depicted in Figure 18 B, reflecting the results of other ssDNA binding assays. These results clearly demonstrate that Yaf does not bind to ssDNA under the tested conditions.

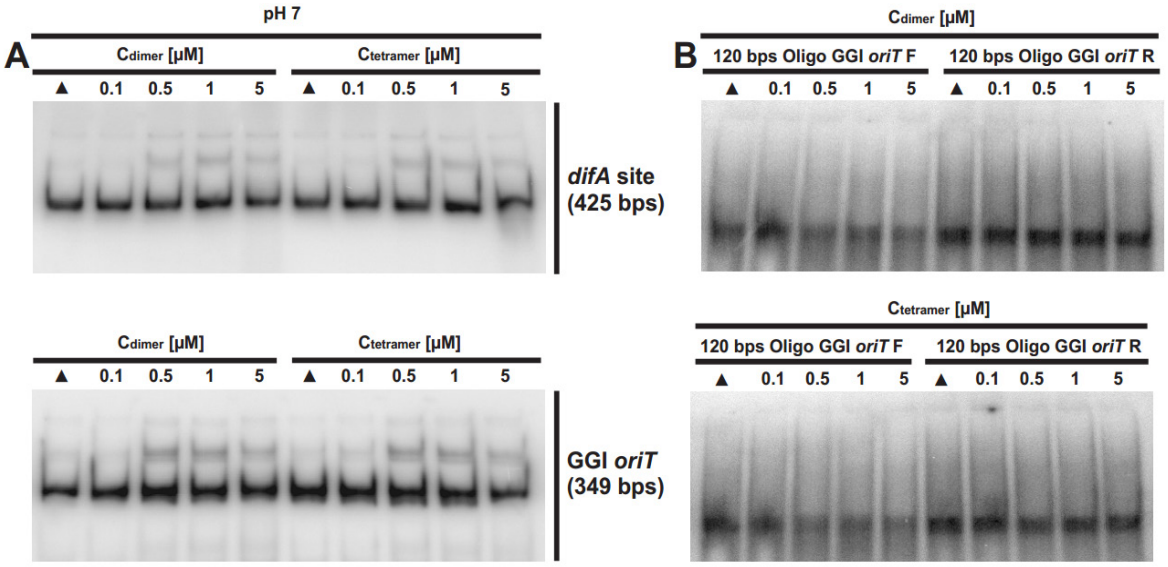


Figure 18: Autoradiographs of DNA binding assays of dimeric and tetrameric Yaf performed with γ-P^{32}-ATP labeled DNA substrates. (A) DNA binding experiment with two dsDNA substrates tested in buffer conditions at pH 7.0 analyzed on 7.5% native TB gels. The lanes are labeled with the corresponding concentrations of dimeric and tetrameric Yaf. The DNA controls are indicated with black triangles. The corresponding dsDNA fragments are indicated on the right site. (B) DNA binding assays performed with two different ssDNA substrates analyzed on 7.5% native TB gels. The corresponding ssDNA substrate is indicated on top. The labeling of the lanes corresponds the labeling described in 'A'.

4.3.8 A Co^{2+} and Mn^{2+} dependent sequence unspecific nuclease activity is associated with dimeric and tetrameric Yaf

To test whether the DNA binding activity of Yaf is influenced by the presence of divalent cations, DNA binding assays were performed in the presence of $Mg^{2+,}$ Mn^{2+} and Co^{2+} and several other possible metal cofactors (Ca^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}). The obtained results revealed a nuclease activity that is associated with elution fractions containing dimeric and tetrameric Yaf in the presence of Co^{2+} and

Mn^{2+} (see Figure 19). This activity was not observed with the other tested divalent cations (data not shown). The degradation of the DNA was observed for different PCR products and was sequence unspecific. Control experiments with heat inactivated Yaf did not show DNA degradation. The degradation of dsDNA leads to the appearance of small nucleotide fragments in the lower part of the gels (labeled with 'b'). Remarkably, the level of nuclease activity observed in fractions containing dimeric and tetrameric Yaf differed between both oligomeric states and was stronger in fractions containing dimeric Yaf.

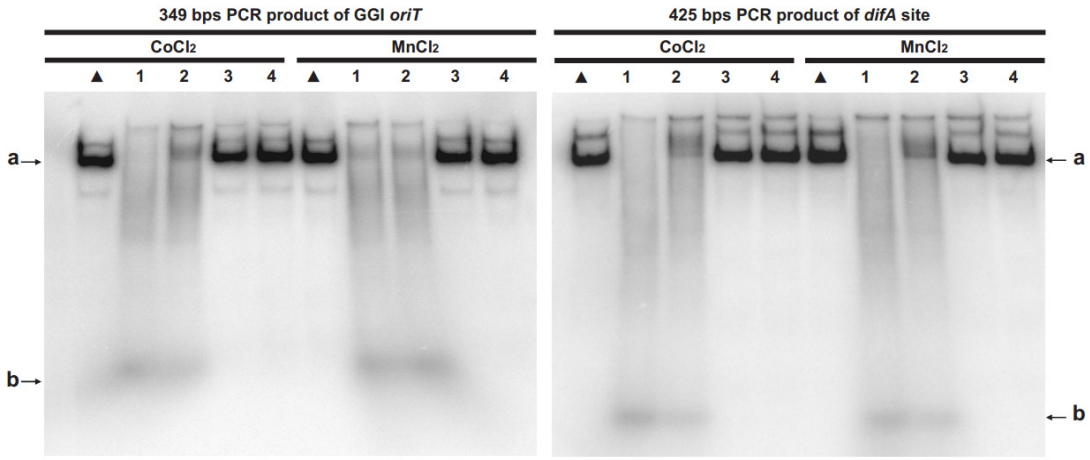


Figure 19: Autoradiographs of DNA degradation assays with His-tagged Yaf and two γ-P^{32}-ATP labeled dsDNA substrates. Left: Native TB PAGE (7.5%) of DNA degradation assay using a 349 bps PCR product with the sequence of GGI *oriT* as dsDNA substrate. The assay was performed with 2 µM of dimeric (1) and tetrameric (2) Yaf in the presence of $CoCl_2$ and $MnCl_2$. Negative controls with inactivated proteins were performed for dimeric (3) and tetrameric (4) Yaf. DNA controls are indicated with black triangles. PCR fragments are indicated with the black arrow (a) and degradation products are indicated with the black arrow (b) on the left site. Right: Native TB PAGE (7.5%) of DNA degradation assay with a 425 bps PCR product containing the *difA* site. The labeling corresponds to the previous description.

To test if Yaf also degrades ssDNA substrates in the presence of different divalent cations, similar experiments were performed with different oligonucleotides. Figure 20 shows an overview of the nuclease activity of Yaf on five different oligonucleotides. Two complimentary oligonucleotides containing the sequence of the GGI *oriT* (see Figure 20 A), as well as two complementary oligonucleotides containing the sequence of the F-plasmid *oriT* (see Figure 20 B) and a T_{74} oligonucleotide (see Figure 20 C) were used as ssDNA substrates. Interestingly, the nuclease activity of Yaf on ssDNA substrates was again only observed in the presence of Mn^{2+} and Co^{2+}, but not in the presence of Mg^{2+}. Both GGI *oriT* containing oligonucleotides and the 'rev' F-plasmid *oriT* containing substrate were completely degraded (see Figure 20 A and B (right gel)). The degradation of the F-plasmid *oriT* 'fwrd' substrate occured via an intermediate fragment ('b') that migrates only slightly faster than the unprocessed substrate (see Figure 20 B (left gel)). Remarkably, the T_{74} substrate was not degraded, regardless of the presence of metal cofactors. Remarkably, the degradation of ssDNA was only observed for elution fractions containing dimeric Yaf and not for tetrameric Yaf.

Taking these results into account, it can be concluded that the Mn^{2+} and Co^{2+} dependent nuclease activity associated with fractions containing dimeric and tetrameric Yaf is sequence unspecific and that dsDNA and ssDNA substrates can be degraded, although the degradation of ssDNA substrates

seems to require the formation of secondary structures. Comparing the degradation efficiency of dimeric and tetrameric Yaf on dsDNA and ssDNA substrates, it is obvious that the dsDNA substrates were degraded faster than the ssDNA substrates and that tetrameric Yaf did not degrade ssDNA substrates, demonstrating that the nuclease activity associated with fractions containing dimeric and tetrameric Yaf preferentially degrades dsDNA.

The DNA degrading activity was only observed in elution fractions containing dimeric and tetrameric Yaf. A background activity caused by a contamination of these fractions can be most likely excluded since no contaminating proteins were detected on Coomassie stained SDS gels and by mass spectrometry analysis, however, a possible contamination with DNAse cannot be completely excluded.

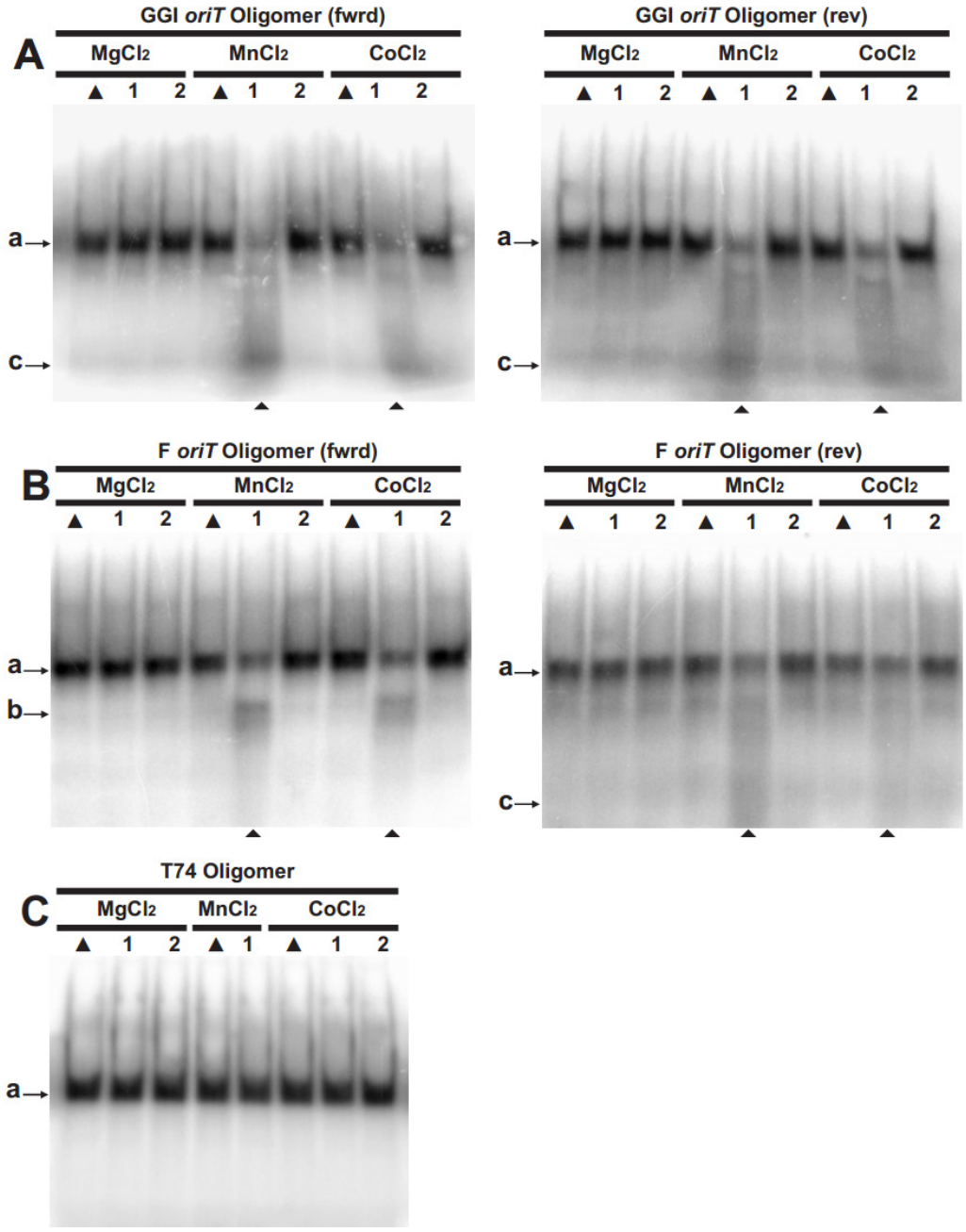


Figure 20: Autoradiographs of DNA degradation assays with His-tagged Yaf dimer and different ssDNA substrates. Samples were analyzed on 7.5% native TB gels. DNA controls are indicated with black triangles. Lanes with active Yaf dimer are labeled '1', lanes with inactivated protein are labeled '2'. All ssDNA substrates were tested with Mg^{2+}, Mn^{2+} and Co^{2+} as metal cofactor. Unprocessed oligonucleotides are indicated with the black arrow (a) on the left site, degraded fragments are indicated with the black arrows (b) and (c). Lanes with obvious degradation activity are indicated with black triangles below the gels. (A) TB PAGEs of degradation assays performed with two complementary 75 bps oligonucleotides with GGI *oriT* sequence. (B) TB PAGEs of degradation assays performed with two complementary 75 bps oligonucleotides with F-plasmid *oriT* sequence. (C) TB PAGE using a T_{74} oligonucleotide as ssDNA substrate.

4.3.9 Yaf does not degrade supercoiled plasmid DNA but shows a sequence unspecific metal dependent relaxation activity enhanced in the presence of Mn^{2+} and Co^{2+}

To characterize the DNA processing activity of Yaf further, DNA relaxation assays were performed with supercoiled plasmid DNA substrates in the presence of different metal cofactors.

All Yaf containing elution fractions from the size exclusion chromatography were tested for relaxation activity. Reactions were performed in the presence of 10 mM Mg^{2+}. Plasmid pEP074 which contains the GGI *oriT* sequence was converted to the relaxed form when incubated with elution fractions containing dimeric and tetrameric forms of Yaf (see Figure 21 A). The same assay was performed using supercoiled pEP073 plasmid DNA containing the F-plasmid *oriT* sequence as substrate. The obtained results were similar to the results observed for plasmid pEP074 (data not shown). Since the relaxation activity only occurs in elution fractions containing dimeric and tetrameric Yaf, it can be concluded that the relaxation activity is associated with elution fractions containing these two oligomeric states of Yaf.

To obtain information on the efficiency of the relaxation reaction of Yaf, a relaxation assay was performed over a time period of 1 h. The results are depicted in Figure 21 B. Samples were taken each 5 min and the reactions were stopped immediately. An initial relaxation of the supercoiled DNA ('II') was already observed after an incubation time of 5 min. Interestingly, a faster migrating reaction intermediate ('III') started to occur after an incubation time of 35 min. At this time point almost all supercoiled DNA was processed to the relaxed form ('I'). The observed fragment ('III') most likely represents a linearized DNA fragment.

Yaf containing fractions were further incubated with relaxed DNA generated by relegation of linearized plasmid ('I'), supercoiled DNA ('II'), and a linearized plasmid ('III') (see Figure 21 C). The results show that Yaf containing fractions only relaxed supercoiled plasmid DNA ('II'), whereas no effect was observed for relaxed ('I') or linearized plasmids ('III'). Since Mg^{2+} was provided as metal cofactor and the previously described degradation activity is dependent on Mn^{2+} and Co^{2+}, no degradation activity on the linearized DNA fragment was observed under the conditions used. The results obtained for the religated plasmid DNA show an additional slower migrating fragment ('IV'). This fragment might represent two concatenated plasmids that migrate with a higher molecular mass. But, since this fragment was also observed in lower intensities in the DNA control and in the negative control with heat inactivated protein, it is most like an artifact of the religation reaction and not a Yaf catalysed reaction. This experiment clearly demonstrates that the relaxation activity associated with dimeric and tetrameric Yaf containing fractions is a metal dependent reaction leading to a convertion of supercoiled plasmids to their relaxed state (see Figure 21 C, lanes 1-4).

Taking the previously observed metal dependency of the DNA degradation activity into account, it was tested if Mn^{2+} and Co^{2+} also influence the relaxation activity of Yaf. A comparison of the results obtained with Mg^{2+}, Mn^{2+} and Co^{2+} is depicted in Figure 22. Both, Mn^{2+} and Co^{2+}, increased the relaxation activity of Yaf. In the presence of Mg^{2+} even the highest Yaf concentrations did not lead to a complete relaxation of the substrate. Thus, the relaxation reaction of Yaf is enhanced by Mn^{2+} and Co^{2+} in comparison with Mg^{2+} and is most efficient in presence of Mn^{2+}.

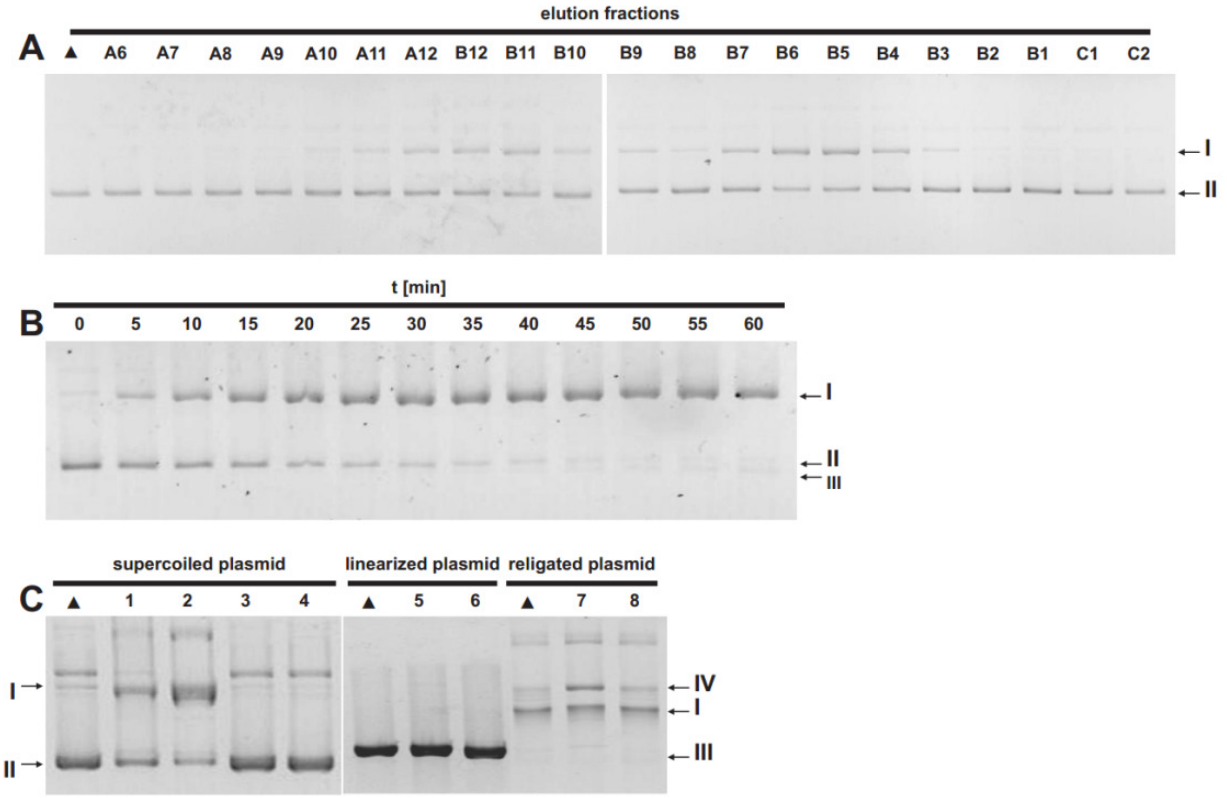


Figure 21: DNA relaxation assays with His-tagged Yaf. Mg^{2+} was provided as metal cofactor in all reactions. (A) DNA relaxation assay performed with Yaf containing elution fractions from gelfiltration. pEP074 was used as supercoiled DNA substrate. The DNA control is indicated with the black triangle. The lanes are labeled corresponding to the elution fractions. (B) Time range experiment of DNA relaxation activity of Yaf performed with supercoiled pEP074 substrate. The samples were taken each 5 min. The lanes are labeled with the corresponding time points. The 0 min sample represents the negative control. (C) Analyses of DNA relaxation and processing activity of Yaf involving three different DNA topologies. The different topologies are indicated on top of the gels. The corresponding DNA controls are indicated with black triangles. The labeling of the lanes corresponds to following reactions: 1 - Yaf incubated 5 min at 95°C after the performed reaction; 2 - reaction with standard conditions; 3 - reaction without Mg^{2+}; 4 - reaction with additional EDTA; 5 - reaction with standard conditions; 6 - reaction with additional EDTA; 7 - reaction with standard conditions; 8 - reaction with additional EDTA. The black arrows at both sites of the gels indicate the topology states as following: I - relaxed state; II - supercoiled state; III - linearized form; IV - putative plasmid dimer

The observed relaxation activity of Yaf on supercoiled DNA substrates has a high similarity to the relaxation activity of *E. coli* Topoisomerase I. Therefore, further experiments were performed to distinguish between both enzymatic activities. It has been reported that Topoisomerase I is strongly dependant on the metal cofactor Mg^{2+} and shows a reduced activity in the presence of Mn^{2+} and Co^{2+} (293). By contrast, the relaxation activity of Yaf was increased in the presence of Mn^{2+} and Co^{2+}, and

showed reduced levels in the presence of Mg^{2+}. Moreover, the incubation of Yaf containing fractions with supercoiled plasmid DNA also resulted in the formation of a linear fragment. These observations were used to distinguish between the enzymatic activities of Yaf and a possible contamination with *E. coli* Topoisomerase I. A comparison of the relaxation activity of both enzymes in the presence of either Mg^{2+}, Mn^{2+} or Co^{2+} is shown in Figure 22.

Topoisomerase I processed the supercoiled plasmid DNA to the relaxed form via the characteristic reaction intermediates forming a "ladder" on the agarose gel (294). Relaxation assays indicated a possibly slightly higher activity of Topoisomerase I in the presence of Mg^{2+}, but did not differ much between the three metal cofactors. Comparing the results obtained from the performed relaxation assays, five differences between both enzymes are observed: I) the activity of Yaf containing fractions was contrary to Topoisomerase I enhanced in the presence of Mn^{2+} and Co^{2+}, II) the relaxation reaction of Yaf containing fractions showed no Topoisomerase I typical reaction intermediates, III) incubation with Yaf containing fractions resulted in a band migrating at the position of the linearized plasmid which is not observed after incubation with Topoisomerase I, IV) the relaxation activity of Yaf co-eluted with dimeric and tetrameric Yaf (35 kDa and 70 kDa), whereas Topoisomerase I elutes with a molecular mass of 97.5 kDa, and V) Topoisomerase I was not detected in fractions containing Yaf at levels necessary for the observed activity. Considering these differences, a background activity of *E. coli* Topoisomerase I resulting from a possible contamination of the recombinant overproduced Yaf can be excluded.

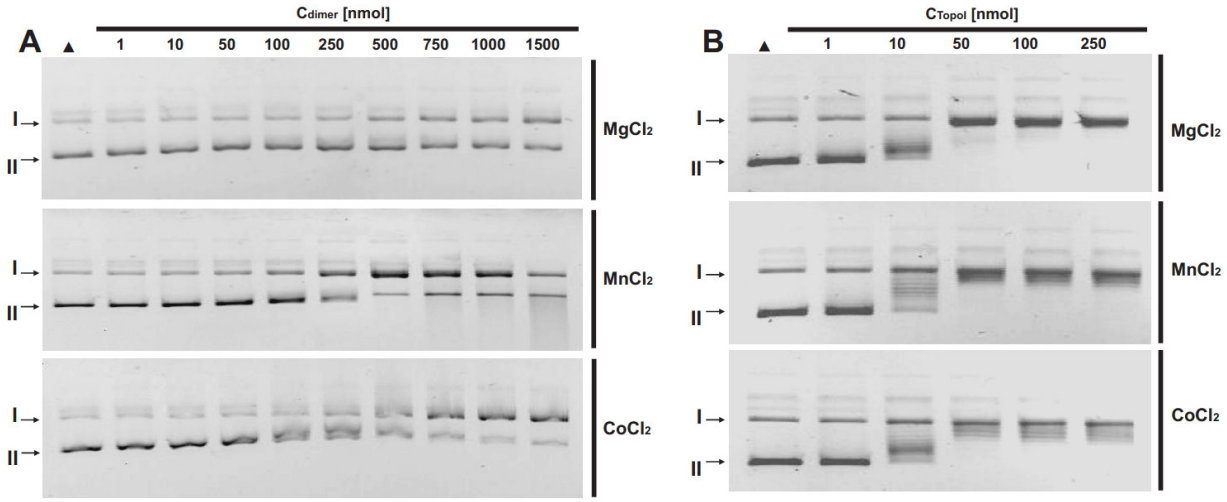


Figure 22: Comparison of DNA relaxation activity of Yaf and *E. coli* Topoisomerase I. The provided metal cofactors are indicated on the right site. The DNA control is indicated with the black triangle. The lanes in 'A' and 'B' are labeled with the corresponding protein concentration. The relaxed form of the DNA is indicated with the black arrow (I) and the supercoiled form is indicated with the black arrow (II) at the left site. Supercoiled pEP074 plasmid DNA was used as substrate. (A) DNA relaxation activity of Yaf. (B) DNA relaxation activity of *E. coli* Topoisomerase I.

4.3.10 Point mutations lead to significant changes in the solubility of Yaf

The identification of the catalytical sites of Yaf is very important to further characterize the enzymatic activity of the protein. Therefore, attempts have been made to identify homologous proteins with

bioinformatical approaches. Such proteins can be used to find conserved residues in the sequence of Yaf, which might form the catalytical site of the protein. Initially, Yaf was only identified in Blast searches in a few *N. gonorrhoeae* strains and no other homologous proteins were found. Only recently, two other homologous proteins could be identified in *N. meningitides* and *N. bacilliformis*. The homology between the *N. bacilliformis* protein and Yaf is very low, but a few conserved residues have been identified, highlighted in light blue in the Alignment in Figure 23. Based on these conserved residues several point mutations were introduced into the sequence of Yaf by site directed mutagenesis as described in Chapter 3.3.6. Besides the conserved residues also mutants from a few other residues of possible importance for the DNA processing reaction were generated. The mutants were overexpressed, purified and tested for there DNA relaxation activity as performed for the wild type protein. A summary of the obtained results is given in Table 52. Remarkably, all single point mutations resulted in the formation of high amounts of inclusion bodies, even at a lower growth temperature of 18°C after induction. Besides the formation of inclusion bodies also the amounts of soluble protein differed strongly from that observed for the wild type constructs. Only small amounts of soluble proteins were obtained and could be used for purifications, resulting in very low yields of the isolated mutants. Before testing the purified mutants for DNA relaxation activity, most of them were concentrated with Amicons. So far, all analyzed mutants showed a metal dependent DNA relaxation activity like the wild type protein. Up to now, the mutants were not tested for the DNA degradation activity.

Taken all these observations together, it is obvious that most of the single point mutations introduced into the sequence of Yaf result in destabilization or misfolding of the protein. The elution profiles of size exclusion chromatographies performed with different mutants also indicated a significant influence of the mutations on the oligomerization of Yaf, which corroborates this assumption. A catalytical site could not be identified up to now, but at least two different sites can be asummed to be involved into the DNA processing reaction of Yaf, which must be a metal ion coordination site and DNA binding or processing site, most likely in close proximity to each other.

Table 52: Overview on expression levels, solubility, purification properties, and enzymatic activity of Yaf mutants

Mutant	Expression	IB fomration	Soluble protein	Purification	Activity
Wild type	high	no	yes	high binding	relaxation
Y47F	medium	yes	low amounts	low binding	relaxation
E57A	medium	yes	low amounts	medium binding	relaxation
D89A	medium	yes	low amounts	low binding	relaxation
D90A	medium	yes	low amounts	low binding	relaxation
D89A/D90A	medium	yes	low amounts	low binding	relaxation
F91A	low	yes	low amounts	low binding	relaxation
Y100F/Y104F	medium	yes	low amounts	low binding	relaxation
S107A	low	yes	low amounts	low binding	relaxation
F115A	low	yes	low amounts	low binding	relaxation
W33A	low	yes	low amounts	low binding	n.a.
E49A	low	yes	low amounts	low binding	n.a.
W105A	low	yes	low amounts	low binding	n.a.
F112A	low	yes	low amounts	low binding	n.a.
DDD-GEA	high	no	yes	high binding	n.a.

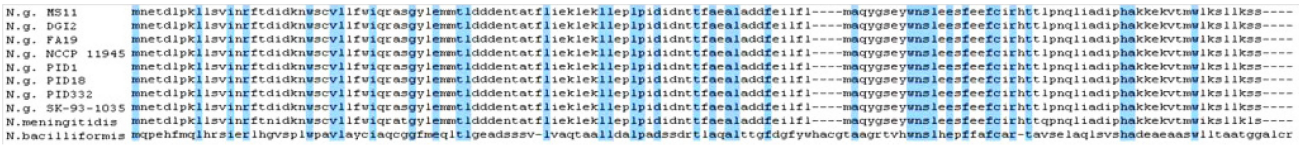


Figure 23: Aligment of Yaf from different *N. gonorrhoeae* strains as well as homologues proteins from *N. meningitides* and *N. bacilliformis*. The overall conserved residues are labeled in light blue.

4.3.11 Crystallization of Yaf

The crystallization of Yaf is performed in close collaboration with the group of Prof. L. Schmitt from the Heinrich Heine University in Düsseldorf.

Before initial crystallization screenings were performed, the oligomeric state of isolated N-terminal His-tagged Yaf was analyzed with dynamic light scattering. These data demonstrated that His-tagged Yaf indeed forms a stable tetramer. The fractions with tetrameric protein were used to screen the initial crystallization conditions. The optimal growth conditions for crystals of Yaf with N-terminal His-tag have been identified with 100 mM Mops/NaOH pH 8.0 (or 100 mM Tris/HCl pH 9.0), 100-250 mM Mg-Acetate and 16% (w/v) PEG 8000. An overview with several examples on the growth of Yaf crystals is given in Figure 24. Interestingly, native Yaf did not crystallize under the conditions optimized for Yaf with N-terminal His-tag and up to now, no growth of crystals with the native protein could be observed. This suggests that the Yaf crystallizes better in its tetrameric form.

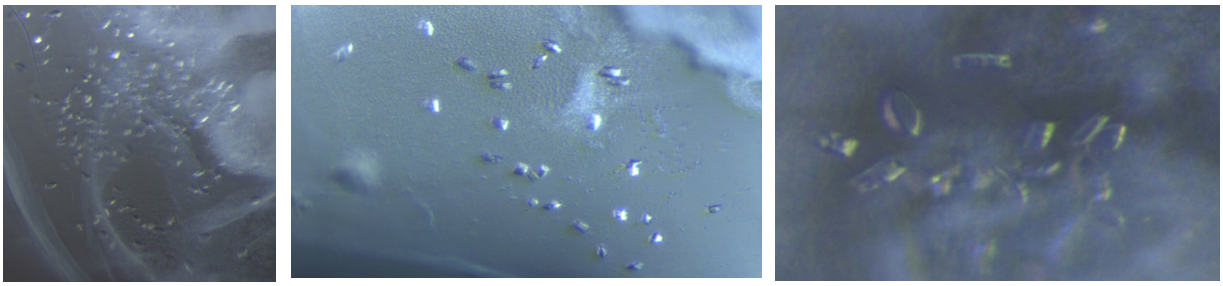


Figure 24: Examples of protein crystals obtained with crystallization conditions using tetrameric N-terminal His-tagged Yaf and buffering conditions of 100 mM Mops/NaOH pH 8.0 (or Tris/HCL pH 9.0), 100-250 mM Mg-Acetate and 16% (w/v) PEG 8000. The size and shape of the obtained Yaf crystals differ significantly within the same and different samples.

The X-ray diffraction pattern of the obtained crystals was measured in a Synchrotron beam-line and revealed a crystal packing according to a monoclinic C2 space group. The crystals showed diffraction to a resolution of 3 Å. Figure 25 shows the X-ray diffraction pattern of a Yaf crystal diffracting within a range of 50 to 3.7 Å.

To solve the crystal structure without any structural homologue, phasing information is essential. Therefore, N-terminal His-tagged Yaf was overproduced in minimal medium in the presence of

selenomethionine. This N-terminal His-tagged Yaf was isolated in similar yields and purity as the wild type protein with N-terminal His-tag. Crystals were obtained under the same conditions used for the protein grown in the absence of SeMe. Initial diffraction experiments showed that Selenomethionine is incorporated into the protein, and as soon as a novel dataset of SeMet-Yaf is obtained, it will be tried to solve the structure of Yaf. This will provide structural insight into the remarkable properties of this protein.

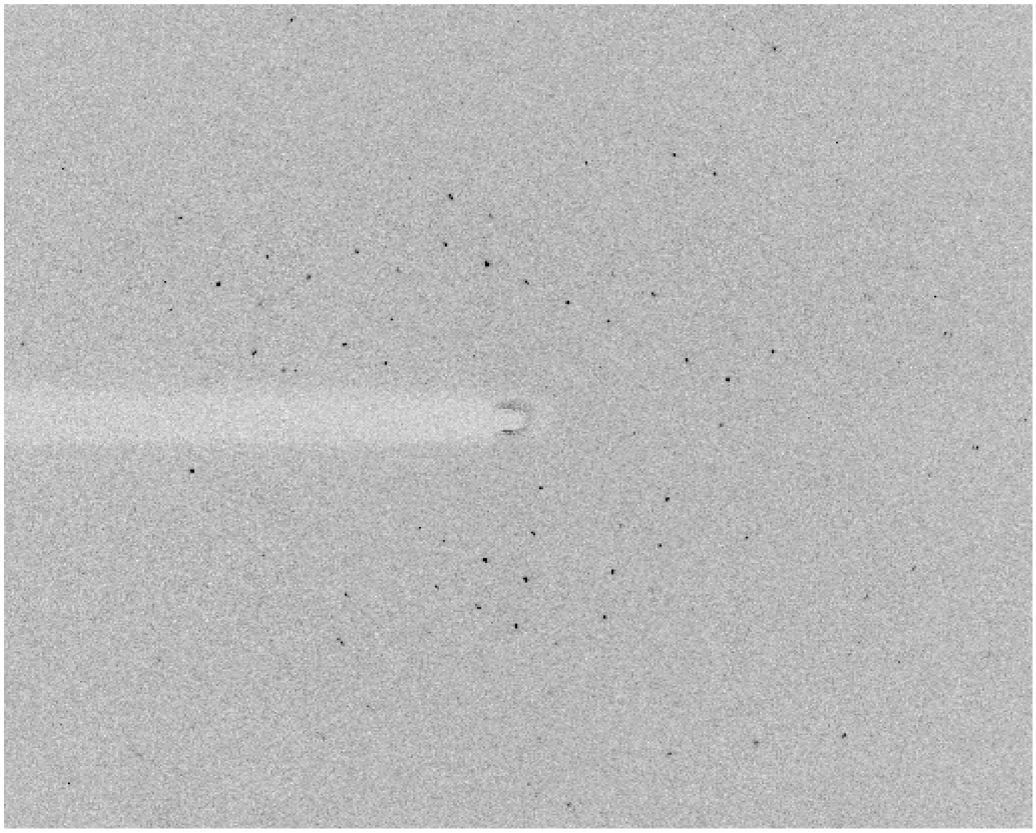


Figure 25: **X-ray diffraction measurement of a crystal with N-terminal His-tagged Yaf. The resolution range of the diffraction patter is within 50 to 3.7 Å. The space group of the crystal packing is C2. The cell dimensions (a b c [Å] α β γ [°]) are measured with 40.194.1177.9 90.089.690.0. The R factor of the highest resolution shell is 36.2%.**

4.4 Biochemical characterization of the relaxase TraI

The relaxase TraI of the novel T4SS of *N. gonorrhoeae* represents the prototype of the uncharacterized relaxase family MOB_H (110). Relaxases are essential proteins within conjugative DNA transfer since they are processing dsDNA into single stranded T-DNA at the *oriT* region. Up to now, only little is known about the biochemical properties of TraI and its role in the DNA processing mechanism. TraI consists of at least two different domains, a relaxase domain and a C-terminal DUF1528 domain. At its N-terminus TraI contains a hydrophobic amphipathic α-helix that most likely facilitates an interaction with the inner membrane. The relaxase domain of TraI strongly differs from previously characterized relaxase domains of relaxases from other MOB families and contains several

unique sequence motifs that are described in detail in Chapter 1.8. The function of the DUF1528 domain is still unknown, but this domain is always found fused to a relaxase domain.

The results described below represent the first biochemical characterization of the MOB_H prototype relaxase TraI and shed light on DNA processing mechanism of this novel relaxase family.

4.4.1 Soluble TraI can be obtained with recombinant overproduction in *E. coli*

To study the DNA processing reaction of the neisserial relaxase TraI, it was tried to overproduce the protein recombinantly with a T7 based expression system in *E. coli*. This turned out to be a challenging task, since the expression levels of *traI* were often very low and the overproduction almost always resulted in the formation of inclusion bodies with insoluble protein. For this reason several attempts using different constructs, changing the growth conditions or the expression strains were made to improve the expression level and the solubility of TraI. Many of the different tested conditions were similar to the conditions that have been tested for TraD. A summary of the most promising constructs is given in Table 53.

Interestingly, the location of a tag on either the C-terminal or the N-terminal site of the protein strongly influenced the solubility and the stability of TraI. This observation was made while comparing the expression levels, the solubility and the purification results obtained from overexpressed *traI* constructs with C-terminal His-tag or Strep-tag and constructs expressing *traI* with an N-terminal His-tag. The expression levels of both N-terminal tagged constructs, as well as for the C-terminal Strep-tagged construct were significantly less compared to the expression level of the construct producing C-terminal His-tagged TraI. Both constructs expressing C-terminal tagged *traI* led to a high production of inclusion bodies and only very low amounts of soluble protein could be obtained after cell breakage. However, the expression of the two constructs producing N-terminal tagged TraI did not result in the formation of inclusion bodies and soluble TraI was observed after cell breakage. Considering these results, all further generated constructs were designed to contain either N-terminal tags or fusion-genes to enhance and optimize the yield of soluble TraI. With most of these constructs higher yields of expression and significantly reduced amounts of inclusion bodies were observed. Two truncated constructs, the relaxase domain of TraI with an N-terminal eXact-tag and the middle domain of TraI with an N-terminal His-MBP-TEV fusion, resulted in the production of insoluble proteins after overexpression, which could not be used for further isolation approaches. The fusion of TraI with a cytosolic MBP was initially thought to improve the solubility and stability of the protein during the folding process and in the following purification. Both MBP fusion constructs indeed resulted in very high amounts of soluble protein, but, like almost all of the other constructs the protein tended to aggregate during further isolation steps, which will be discussed in more detail in the following Chapter. Only one construct which is based on a vector that is normally used to fuse an N-terminal eXact-tag to the target protein finally resulted in the production of soluble TraI in such quatities that TraI could be purified. Initial overexpression experiments with this constructs resulted in high amounts of soluble TraI. Remarkably, the overproduced protein did not contain the eXact-tag, but mass spectrometric peptide finger print analysis confirmed the presence of the start methionine within the isolated protein. Sequencing of the construct revealed a frame shift mutation between the multiple cloning site and the inserted *traI* gene. The precise reason for the observed successful overproduction of TraI based on the expression of this ‘frame shift’ construct is up to now not

known, but, most likely the *traI* gene is translated from an internal start-site, resulting in the overproduction of full-length TraI without the N-terminal eXact-tag.

In addition to the numerous tested expression constructs, it was tried to improve the isolation of soluble TraI by using different expression strains. Therefore, the *E. coli* strains BL21(DE3), BL21(DE3) Star and LEMO21(DE3) were tested in various growth and expression conditions. The most successful overproduction of native TraI could be observed using BL21(DE3) Star cells grown in LB medium incubated at 37°C in the presence of 0.5% glucose. Remarkably, a lowering of the expression temperature also resulted in lower expression levels of *traI*.

Table 53: Summary of overproduction and purification results obtained from different *traI* constructs

Protein	Protein amount	Protein soluble	Inclusion bodies	Purification successful
TraI C-terminal His-tag	medium	in part	yes	no
TraI C-terminal Strep-tag	low	in part	Yes	no
TraI N-terminal His-tag	low	yes	No	in part (aggregation)
TraI N-terminal Strep-tag	low	yes	No	in part (aggregation)
TraI N-terminal eXact-tag	high	Yes	No	no
TraI N-terminal MBP fusion	high	Yes	no	in part (aggregation)
TraI N-His-MBP-TEV fusion	very high	Yes	no	in part (aggregation)
TraI native	high	Yes	no	yes (soluble protein)
TraI (relaxase dom.) N-terminal eXact-tag	high	No	yes	no
TraI (middle dom.) N-His-MBP-TEV fusion	medium	No	no	no
TraI (DUF1528 dom.) N-terminal His-tag	high	Yes	no	in part (aggregation)

4.4.2 Soluble native TraI can be isolated in the presence of the detergent DDM

The isolation of purified TraI was a challenging and difficult task since TraI strongly tends to aggregate, either already during overproduction or during the purification. Many different approaches with different constructs, overexpression conditions and purification protocols were performed to obtain soluble TraI. Initial attempts to purify TraI containing a His- or Strep-tag resulted in very low amounts of TraI that still contained a high number of contaminats. Mass spectrometry analysis indicated that many of these contaminations were chaperones, which suggests a misfolding of TraI. Indeed, size exclusion chromatography showed that these proteins eluted in the exclusion volume of the column.

Two constructs expressing *malE-traI* fusions have been generated to prevent protein aggregation and to obtain higher protein yields. The fusionprotein without a His-tag and a TEV cleavage site was overproduced in high yields and was purified in one step using an MBPTrap (GE Healthcare) column. Unfortunately, the MBP-TraI fusionprotein only had a weak binding affinity to the column material and only very low amounts of the protein were eluted. Most of the protein remained unbound in the flowthrough. The other MBP-TraI fusionprotein with an N-terminal His-tag and a TEV cleavage site

between both proteins yielded in high amounts of soluble protein and was purified using two affinity chromatography steps followed by a size exclusion chromatography. The final amounts of purified protein were much higher than the obtained yields from the other MBP-TraI fusionprotein. Nevertheless, like in previous purification attempts also this purification ended up with the elution of TraI in the exclusion volume of the gelfiltration column, indicating aggregation of the protein. These results represented for a long time the best purification of TraI. Unfortunately, attempts to cleave off the MBP protein by TEV protease and to further purify the native TraI were not successful.

Besides different approaches to purify tagged proteins, attempts have been made to isolate native TraI by a combination of different purification methods like ion exchange chromatography, hydrophobic interaction chromatography, and size exclusion chromatography. Unfortunately, this purification protocol including a Q-Sepharose column, a Phenyl-Sepahrose column, a Hydroxyapatite column, and a Superdex 200 gelfiltration column, resulted in >95% loss of the protein during the different purification steps and the purified TraI eluted again as an aggretated protein in the final gel filtration step. The obtained protein yields from the purification of native TraI were slightly lower than the obtained yields in the purification of the His-MBP-TEV-TraI.

An important breakthrough in the purification of soluble native TraI was made by performing an extensive screen of different detergents. The addition of the detergent DDM to the different purification steps prevented aggragation of TraI during the different purification steps, and resulted in the purification of soluble and active TraI. The results of the purification are shown in Figure 26. In the performed SDS-PAGE analysis using 11% gels, TraI migrates at a position corresponding to a molecular mass of 120 kDa.

The successful purification protocol includes an additional solubilization step with DDM after cell disruption. The further purification of the protein includes an anion exchange chromatography step followed by a cation exchange chromatography step (see Figure 26 A) and a second cation exchange chromatography step to concentrate the protein. The final purification step is performed with size exclusion chromatography using a Superdex 200 column (see Figure 26 B). In the first anion exchange chromatography, TraI did not bind to the column material under the conditions used and remained in the flowthrough. Most of the contaminating proteins were bound to the anion exchange column and, therefore, this step resulted in a significant purification of TraI. The flowthrough was loaded on a cation exchange column and bound TraI eluted at an approximate NaCl concentration of 125 mM (see Figure 26 A (left); Fractions A2 to A4). Since TraI tends to aggregate during concentration with centrifugal filter units, the TraI containing fractions were first diluted and then concentrated in a second cation exchange chromatography step by using a steep elution gradient (see Figure 26 A (middle)). The concentrated elution samples were analyzed on 11% SDS gels (see Figure 26 A (right)). In the final purification step performed by gel filtration most of the contaminating proteins that were still present after the second cation exchange chromatography were removed (see Figure 26 B). Using this purification protocol TraI did no longer elute in the exclusion volume of the gelfiltration column, but eluted at a volume corresponding to a molecular mass of approximately 280 kDa, suggesting the existence of a higher oligomeric state of TraI, which will be discussed in detail in the following Chapter. The TraI containing fractions still contain some contaminating proteins which could not be removed. These proteins were analyzed by mass spectrometry and most of them were identified as degradation products of TraI. With this purification protocol approximately 0.5 mg of TraI can be isolated per 1 L *E.coli* overexpression culture.

Relaxases are difficult proteins to purify and a successfull isolation of the full-length protein has been reported for only a few of them. The hereby described purification method is the first successful protocol to isolate a native soluble full-length relaxase of the novel and uncharacterized MOB_H family.

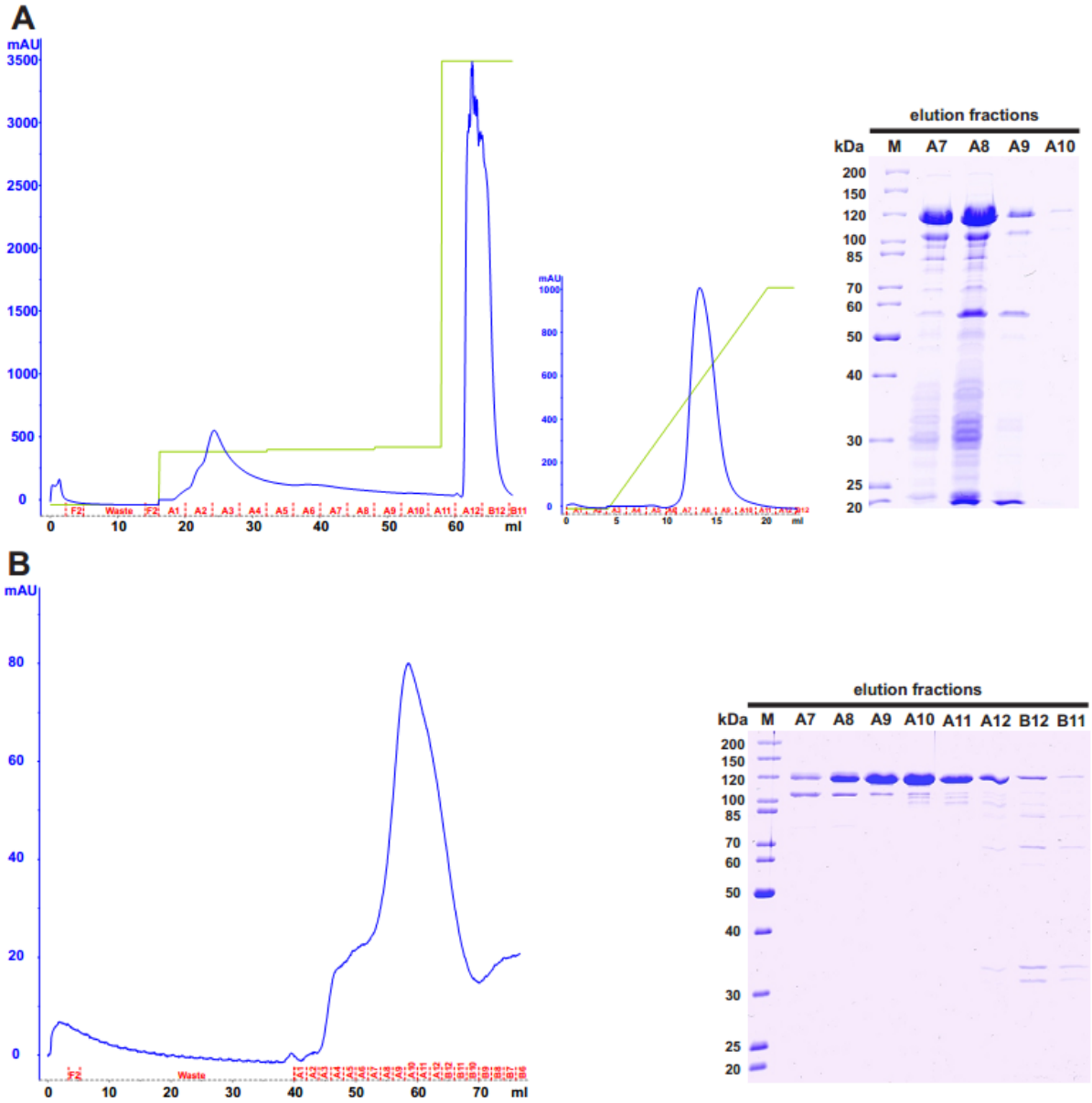


Figure 26: Isolation of native TraI in the presence of DDM. (A) Elution profiles of the first cation exchange chromatography step (left) and the second cation exchange chromatography step (right). Protein was detected with a wavelength of 280 nm (blue line). The intensity of the UV light is given as mAU. The collected fractions are labeled in red. The concentration of the 'Ion exchange TraI detergent B' buffer is indicated as green line (gradient from 0-100%). The elution fractions of the cation exchange chromatography were analyzed with 11% SDS gels (right). Molecular masses are indicated on the left site (PageRuler unstained protein ladder; Fermentas). The labeling of the lanes corresponds the labeling of the elution fractions. (B) Elution profile of size exclusion chromatography (left) and 11% SDS-PAGE analysis of the protein containing elution fraction. The labeling corresponds to the description given in (A).

4.4.3 The oligomeric state of TraI

Purified TraI elutes from the gelfiltration column at a volume which corresponds a molecular mass of approximately 280 kDa. Interestingly, TraI has a calculated molecular mass of 94.7 kDa, but migrates on SDS-PAGE gels at a position corresponding to a molecular mass of 120 kDa. Western blot analyses using a TraI antibody on *N. gonorohoeae* cells have shown that also the native protein in *N. gonorrhoeae* migrates on SDS-PAGE gels at a molecular mass of 120 kDa (202). The reason for the observed difference of 25 kDa is up to now not known. This difference was also observed in preparations without DDM, excluding an influence of the detergent DDM in this abreviant migration behavior on SDS-PAGE gels. Also other purified TraI proteins expressed from different constructs migrate on SDS-PAGE gels with a higher molecular mass. The abreviant migration behavior of TraI is possibly caused by its hydrophobic N-terminus. Alternatively, TraI does not unfold completely in the presence of SDS or the protein is post-translationally modified.

The observed molecular mass of TraI (280 kDa) on the gelfiltration column seems to fit with the calculated mass of a trimer of TraI (285 kDa). However, the abreviant migration behavior on SDS-PAGE gels might also occur on the gelfiltration column, and the protein might be a dimer (190 kDa) surrounded by a DDM micelle (70 kDa) (295). To further study the oligomeric state of TraI, the isolated protein was analyzed on a Blue Native gel (5-13% acrylamide gradient). The results from this experiment are shown in Figure 27. Six different marker proteins with known masses were used as references to enable an approximate mass estimation of TraI. The molecular masses and the relative mobility of the proteins were plotted as a function of each other (see Figure 27 B). Using this graph to calculate the molecular mass of TraI, the Protein migrated at a mass corresponding to 260 kDa, which is smaller than the calculated mass for a trimer. Taking the abreviant behavior of TraI on SDS-PAGE gels into account, it can be concluded that TraI is purified as an oligomer that is most likely a dimer.

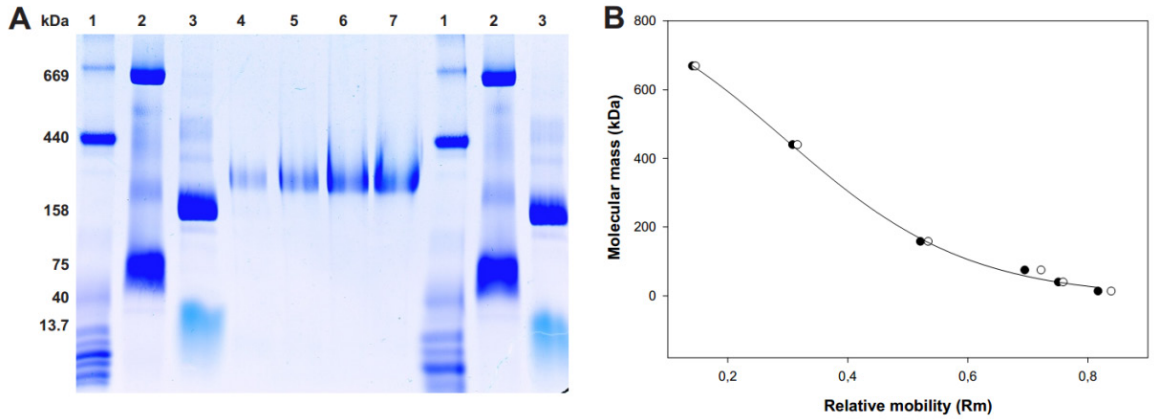


Figure 27 : BN-PAGE analysis of the oligimomeric state of TraI. (A) BN-PAGE with 5-13% acrylamide gradient gel. Molecular masses are indicated on the left site. The labeling of the lanes correspond the following: 1 - Ferritin + Ovalbumine standard; 2 - Thyroglobuline + Conalbumin standard; 3 - Aldolase + Ribonuclease standard; 4 - 458 ng TraI; 5 - 915 ng TraI; 6 - 1.4 mg TraI; 7 – 1.8 mg TraI (B) Plot of the molecular masses of the marker proteins in relation to their relative mobility within BN-PAGE analysis. Each marker protein was loaded and measured in dublicate (black and colorless dots). The trendline correspond a sigmoidal regression.

4.4.4 Limited proteolysis elucidates the domain organization of TraI

Analysis of the primary amino acid sequence of TraI suggested that TraI consists of different domains including an N-terminal hydrophobic amphipathic α-helix (4-5 kDa), followed by a relaxase domain (42.5 kDa), a non conserved domain of unknown function (34 kDa) and a DUF1528 domain (11.8 kDa). To obtain more information about the domain structure of TraI, limited proteolysis with trypsin and ProteinaseK was performed. Therefore, TraI was incubated with increasing concentrations of trypsin and ProteinaseK and the samples were analyzed on 13-15% gradient SDS gels.

Interestingly, TraI showed significant diffences in the sensitivity for both proteases. Comparing the concentrations of trypsin and ProteinaseK which are necessary to initiate the proteolysis of TraI (1 nM for trypsin vs. 0.04 nM for ProteinaseK) and to reach a complete digest of the protein (12 nM for trypsin vs. 1.28 nM for ProteinaseK), it is obvious that TraI is very sensitive to ProteinaseK and is less sensitive to trypsin. Both proteolysis experiments do not only differ with regard to their protease sensitivity, but also with respect to the migration pattern of the obtained fragments. Allmost all observed fragments from the trypsin proteolysis (see Figure 28 A) are also present in the ProteinaseK digest, but some fragments are only observed in the ProteinaseK digestion (see Figure 28 B). Whereas the proteolytic digest with trypsin resulted in the formation of specific fragments, the digest with ProteinaseK seems to result in an initial digestion of TraI into specific fragments followed by further digestion of these fragments.

At least a trypsin concenteration of 1 nM was necessary to detect a proteolytic activity on TraI. Increasing trypsin concentrations finally ended up with seven disting fragments which could be observed within trypsin concentration >1 nM, representing different domains and peptide fragments of TraI. Initial fragments migrating at positions corresponding to molecular masses of 100 (‘2’), 85 (‘3’) and 35 kDa (‘6’) were observed within trypsin concentrations of 1-8 nM. At trypsin concentrations of > 4 nM an additional fragment (‘4’) appeared at the position corresponding to 80 kDa and fragment ‘2’ (100 kDa) was almost complete digested. Next to the occurence of fragment ‘7’ (26 kDa), fragment ‘6’ (35 kDa) started to disappear at trypsin concentrations >8 nM. Finally, the last two fragments ‘5’ (55 kDa) and ‘8’ (17.5 kDa) appeared at trypsin concentrations above 16 nM. Considering the observed masses of the fragments ‘4’, ‘7’ and ‘8’ a total mass of 123.5 kDa will be obtained, corresponding to the mass of undigested TraI. Based on Figure 28, the following model might be conclusive:

Fragment ‘1’:	resents the full-length protein in an undigested state.
Fragment ‘2’:	first formed fragment that is also produced in absence of protease. Most likely it results n degradation of TraI.
Fragments ‘3’ and ‘6’:	derive from fragment ‘2’ and start to appear when fragment ‘2’ disappears. Most likely ment ‘3’ represents the relaxase and the middle domain of TraI and fragment ‘6’ represents DUF1528 domain and a linker region between middle domain and DUF1528 domain.
Fragments ‘4’ and ‘8’:	gment ‘4’ derives from fragment ‘3’ and represents the relaxase domain and/or the middle ain. Fragment ‘8’ derives from fragment ‘7’ and represents the DUF1528 domain.
Fragment ‘7’:	st likely this fragment represents either the DUF1528 domain that is cleaved with an itional sequence part and further digest produce fragment ‘8’, or it represents a part of the dle domain of TraI.

The fragmentation pattern of the proteolytic digest with ProteinaseK revealed a similar domain organization as observed with trypsin and confirmed the assumption that TraI consists of at least three domains. Almost all of the previous described fragments ‘1’ till ‘8’ are also present in the ProteinaseK digested samples. Only fragment ‘5’ was not observed on its expected position, but seems to be represented by the fragments ‘10’ to ‘12’, which migrate in a range of 45-50 kDa. However, in comparison with the obtained results from the trypsin digest the ProteinaseK digest leads to a much stronger fragmentation and degradation of TraI. One of the most obvious differences between both proteolysis experiments is the high sensitivity of TraI against ProteinaseK. This could be especially observed for fragments ‘3’ and ‘4’ that are already almost complete digested at a ProteinaseK concentration of 0.64 nM, whereas within the proteolysis with trypsin both fragments are still present at the highest protease concentration.

These results lead to the conclusion that the N-terminal relaxase domain has a certain resistance against endopeptidases like trypsin, but is very sensitive to exopeptidases like ProteinaseK, which has an endo- and an exopeptidase activity. This difference between both proteases could be a reason for the observed numberous degradation fragments in the ProteinaseK experiment. Most likely, the different sensitivities of TraI to both proteases result from the structural integrity of the observed domains, which possibly influences and restricts the accessibility to endopeptidases. Taking all observed fragments from the proteolytic digest with trypsin into account, it is most likely that TraI constists of at least three different domains.

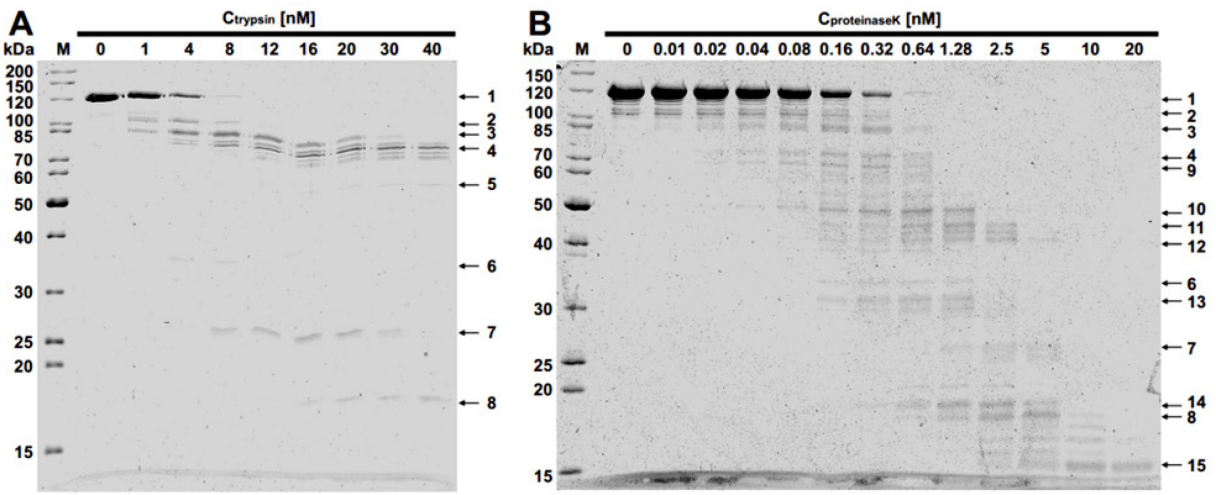


Figure 28: SDS-PAGE analysis of limited proteolysis of TraI. (A) Limited trypsin digest of TraI with increasing trypsin concentrations indicated on top of the lanes. (B) Limited ProteinaseK digest of TraI with increasing ProteinaseK concenterations indicated on top of the lanes. The molecular masses for both gels are given on the left sites (PageRuler unstained protein ladder; Fermentas). The black arrows on the right gel sites indicate the migration positions of TraI (1) and of several fragments obtained from the proteolytic digest (2-15).

4.4.5 TraI binds ssDNA and dsDNA in a sequence unspecific competative reaction

Different DNA binding and competition experiments were performed with native TraI to study the DNA binding mechanism of this novel relaxase family. To characterize the DNA binding reaction of TraI, γ-P^{32} labeled ssDNA substrates were incubated with increasing concentrations of TraI and were analyzed for mobility shifts on 7.5% native TB gels. Besides different substrates it was tested whether different metal cofactors can influence the binding reaction of TraI. It was observed that TraI binds to all tested ssDNA substrates in a sequence unspecific reaction (see Figure 29; comparison of a T_{74} primer with a primer containing the putative *oriT* region of the GGI). The DNA binding reaction was not dependent on a metal cofactor (see Figure 29). The observed DNA binding affinity of TraI was at least 200 nM, however, further experimental approaches like anisotropy measurements with ssDNA substrates are necessary to determine the K_d of the DNA binding reaction of TraI.

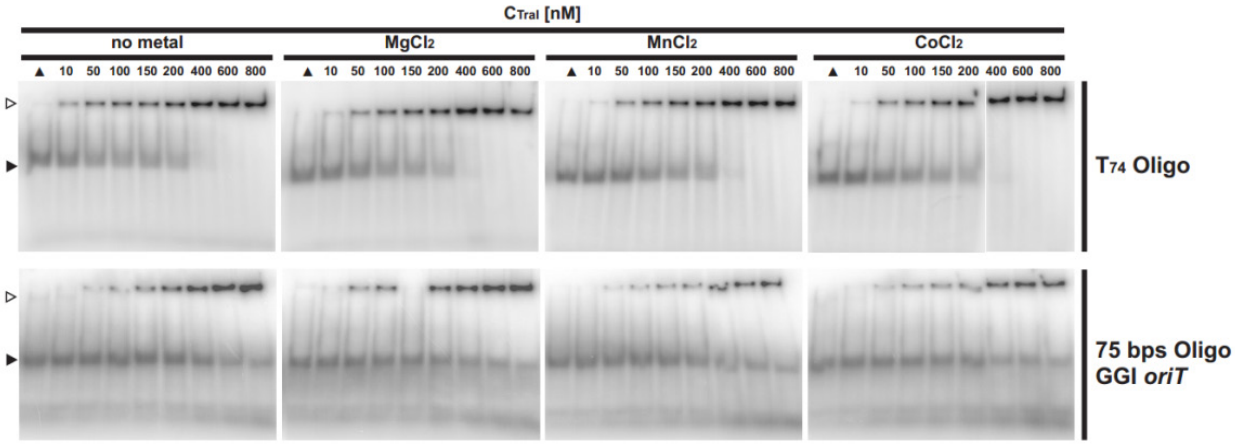


Figure 29: Autoradiographs of DNA binding assays of TraI providing Mg^{2+}, Mn^{2+} or Co^{2+} as metal cofactors and either γ-P^{32} labeled 'T_{74}' oligonucleotide or '75 bps GGI *oriT*' oligonucleotide as DNA substrates. The samples were analyzed on 7.5% native TB gels. Negative controls without TraI are indicated with black triangles. The labeling of the lanes corresponds to the used TraI concentrations (10 - 800 nM). The DNA substrate concentration used in each reaction was 100 nM. The different metals are indicated on top of the gels, the DNA substrates are indicated on the right sites. The migration position of unbound DNA is indicated with black triangles on the left sites, the position of protein-bound DNA is indicated with colorless triangles.

To further characterize the DNA binding reaction of TraI, competition assays with γ-P^{32} labeled dsDNA and ssDNA substrates, which were competed with unlabeled T_n oligonucleotides were performed. At first, two labeled dsDNA substrates either containing the GGI *oriT* sequence or the *difA* site sequence were competed with T_n oligonucleotides in a length range of 6-28 nucleotides (see Figure 30 A and B). Subsequently, a polyT primer with a length of 25 nucleotides and a primer corresponding to the putative *oriT* sequence with a length of 75 nucleotides were competed with T_n oligonucleotides in a length range of 6-28 nucleotides (see Figure 30 C and D).

Interestingly, it could be observed that TraI binds sequence unspecific to all tested DNA substrates, but the DNA binding affinity differs between dsDNA substrates and ssDNA substrates and seems to be higher then 200 nM. Both dsDNA and ssDNA substrate show similar results for the competition of T_n oligonucleotides. However, oligonucleotides shorter than T_9 cannot compete with the binding of

the other DNA substrate, whereas T_n oligonucleotides with a length between 12 and 21 nucleotides competed with increasing efficiency. T_n oligonucleotides that are longer than 21 nucleotides were able to compete completely with the binding of both ssDNA and the dsDNA substrates.

Taking into account that oligonucleotides shorter than 9 nucleotides were not able to compete longer DNA substrates, but T_n oligonucleotides longer than 9 nucleotides were able to compete with the bound DNA substrates, it can be concluded that TraI has a minimal binding frame of 9 nucleotides. The observation that T_n oligonucleotides with increasing lengths are more efficient competitors than shorter onligonucleotides indicates that the binding affinity of TraI for ssDNA rises with increasing length of the oligonucleotide.

Thus, from the obtained results it can be concluded that TraI binds sequence unspecific to ssDNA and dsDNA substrates in a metal independent reaction. The minimal binding frame of TraI for ssDNA substrates was determined to be 9 nucleotides. Increasing the length of the oligonucleotide further increases its binding affinity to TraI. The sequence specificity of the DNA binding reaction of TraI is not influenced by Yaf (data not shown).

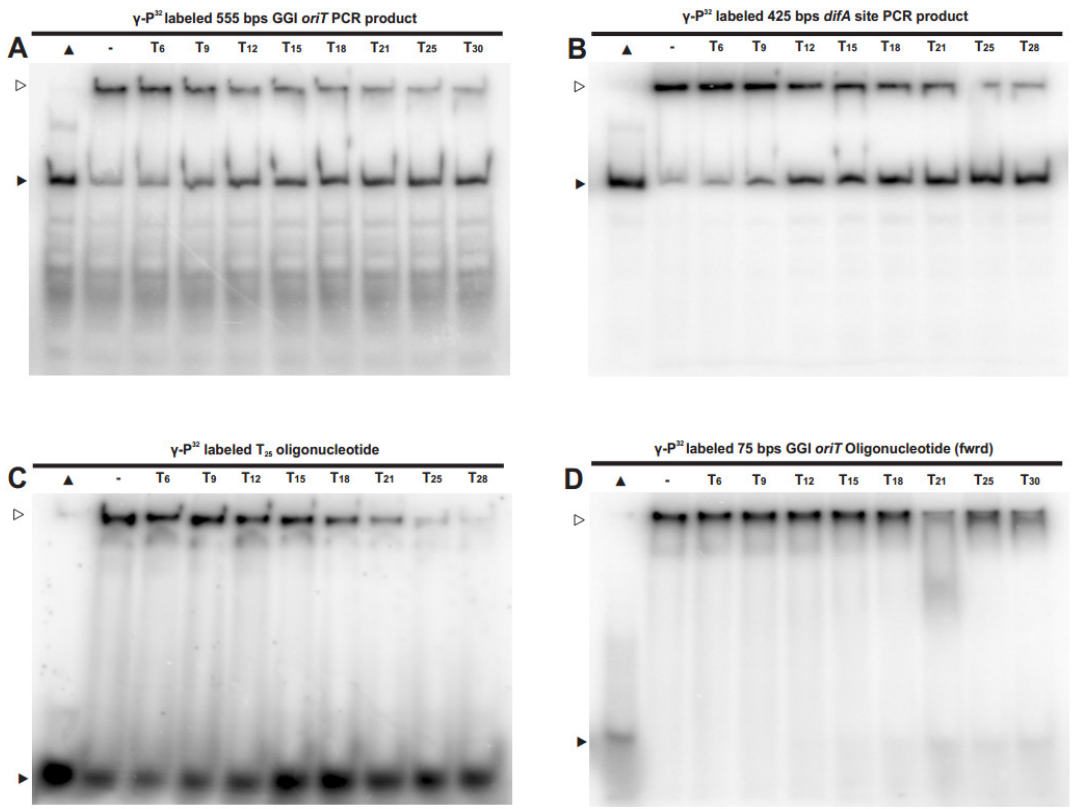


Figure 30: Autoradiography of DNA binding and competition assay with different γ-P^{32}-labeled DNA substrates competed with different unlabeled T_n oligonucleotides. The assay was performed with isolated native TraI. Negative controls without protein are indicated with black triangles and reactions without competitor DNA are indicated with a dash on top of the lanes. The labeling of the lanes indicates the corresponding competitative T_n oligo-nucleotide. The migration position of unbound DNA is indicated with black triangles on the left sites, the position of protein-bound DNA is indicated with colorless triangles. (A) reaction with γ-P^{32}-labeled 555 bps PCR product with GGI *oriT* sequence; (B) reaction with γ-P^{32}-labeled 425 bps PCR product containing the *difA* site sequence; (C) reaction with γ-P^{32}-labeled T_{25} oligonucleotide; (D) reaction with with γ-P^{32}-labeled 75 bps GGI *oriT* oligonucleotide.

4.4.6 TraI relaxes supercoiled plasmid DNA in a metal dependent sequence unspecific reaction similar to the activity of *E. coli* Topoisomerase I

To further characterize the mechanism of the relaxase TraI, DNA relaxation assays were performed as described in Chapter 3.8.6. Several different conditions including various DNA substrates, different divalent cations and changes in the reaction conditions were tested to study the DNA relaxation activity of TraI.

The relaxation activity of TraI was tested on supercoiled DNA substrates containing either the *oriT* sequence from the GGI or from the F-plasmid. Remarkably, both supercoiled DNA substrates were converted from the supercoiled to the relaxed form. Figure 31 A shows the time dependent relaxation activity on a plasmid that contains the *oriT* sequence of the GGI, but a similar result was obtained for the plasmid that contains the *oriT* of the F-plasmid (data not shown). Relaxation assays on plasmids that contain neither the *oriT* sequence from the GGI nor the *oriT* sequence of the F-plasmid, showed that also these plasmids were converted to the relaxed form (data not shown). This suggested that a sequence unspecific relaxation activity was present in the fractions that contained purified TraI. This was an unexpected result since almost all characterized relaxases show a sequence specific relaxation reaction and might indicate that TraI requires an accessory protein to promote efficient cleavage of the *oriT* sequence.

To demonstrate that the slower migration of the different plasmids is caused by relaxation of the supercoiled DNA and not by binding of TraI to the DNA, several experiments were performed. At first, TraI was incubated with supercoiled, linearized and religated plasmid DNA (see Figure 31 B). In this experiment a mobility shift was only observed for the supercoiled plasmid. Secondly, after incubation with TraI, the plasmid was treated with ProteinaseK to degrade the protein. After incubation with ProteinaseK, the slower migrating band was still observed (data not shown). The shift to a band with a slower mobility is dependent on the presence of divalent cations, whereas binding of TraI to the DNA is independent of divalent cations.

In view of the high similarity of the relaxation reaction observed for purified TraI to the enzymatic activity of Topoisomerase I from *E. coli,* further experiments were performed to exclude that the observed reaction is a result of a Topoisomerase I contamination. The final gelfiltration chromatography step of the purification of TraI should separate TraI (elutes at a volume corresponding to 280 kDa) and Topoisomerase I (should elute at 97 kDa). Titration experiments with commericially obtained Topoisomerase I (*New England Biolabs*) showed that the amounts required for the observed relaxation activity should be visible on SDS-PAGE gels and thus should also be easily detectable with mass spectrometry. Topoisomerase I could however never be detected in fractions containing TraI using mass spectrometry (data not shown). In a second experiment, similar to that performed for Yaf (see Figures 22 B and 31 C), the metal dependency and the formation of different intermediates by TraI and Topoisomerase I were compared to distinguish between both activities. Like for Yaf, also the relaxation activity of TraI was observed in the presence of Mg^{2+} but is strongly enhanced in presence of Mn^{2+} and Co^{2+} (see Figure 31 C). Also no typical Topoisomerase I reaction intermediates were observed (see Figure 22 B), but one fragment migrating at the position of the linearized plasmid was observed in the presence of TraI. Considering these differences, it can be concluded that the relaxation activity associated with the fractions containing TraI is not a background activity of *E. coli* Topoisomerase I.

Remarkably, also TraI of the *E. coli* F-plasmid cleaves DNA more efficient in the presence of Mn^{2+} then in the presence of Mg^{2+} (246, 296).

Thus, the fractions containing purified TraI kept a relaxase activity which resembled the activity observed for purified Yaf. Unfortunately, both relaxation activities were sequence independent. TraI showed no strong preference for any of the tested plasmids and those plasmids that contained the putative *oriT* encoded within the GGI were not processed faster. The addition of Yaf to the relaxation assays did also not result in a more specific relaxation activity (data not shown). Purified TraI binds in electrophoretic mobility assays to DNA with a high affinity, demonstrating that the purified TraI is active in DNA binding under the tested conditions, making it unlikely that the inability to detect sequence specific relaxation is caused by inactive TraI. Possibly the relaxation assays were not sensitive enough to detect the specificity of TraI. Therefore, also DNA cleavage assays were performed which will be described in detail in the following Chapter.

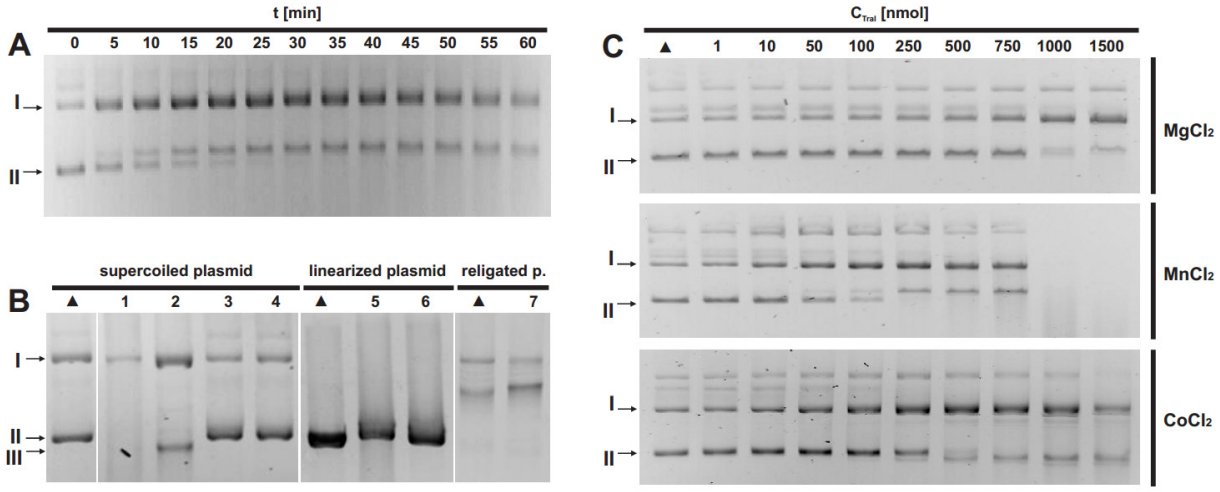


Figure 31: Comparative analyses of the DNA relaxation activities of TraI. (A) Relaxation time range experiment with TraI. The control (T_0) is indicated as black triangle. Samples were taken in 5 min steps and were immediately inactivated with SDS and EDTA. The labeling of the lanes indicates the corresponding reaction time. (B) Influence of the DNA topology on the relaxation activity of TraI. Supercoiled, linearized, and religated DNA derived from plasmid pEP074 was used as substrate. The labeling of the controls and the DNA topolgies 'I' and 'II' correspond to 'A'. The migration position of linearized DNA is indicated with (III). The labeling of the lanes correspond the following: 1 - TraI incubated 5 min at 95°C after reaction; 2 - standard conditions; 3 - no Mg^{2+}; 4 - additional EDTA; 5 - standard conditions; 6 - additional EDTA; 7 - standard conditions (C) Influence of different metal cofactors on the relaxation activity of TraI. The relaxation assay was performed with increasing protein concentrations in presence of Mg^{2+}, Mn^{2+} and Co^{2+} (indicated on the right site). Negative controls without protein are indicated with black triangles on top of the lanes. The migration positions of supercoiled (II) and relaxed (I) DNA fragments are indicated with the black arrows on the left site.

4.4.7 TraI cleaves ssDNA sequence unspecific in presence of Mn^{2+} and Co^{2+}

To test whether TraI cleaves ssDNA, cleavage assays were performed with isolated TraI under many different conditions using several different oligonuceolotide substrates. The experiments were performed as described in Chapter 3.8.7. Five different oligonucleotides have been tested: two complementary oligonucleotides containing the putative *oriT* of the GGI (see Figure 32 A), two

complementary oligonucleotides containing the F-plasmid *oriT* (see Figure 32 B) and a T_{74} oligonucleotide (see Figure 32 C).

The cleavage assays were performed in the presence of Mg^{2+} or Mn^{2+}, and in the presence and absence of ATP and Yaf. Remarkably, cleavage was reproducibly observed for the "GGI *oriT* 'fwrd'" and the "F oriT 'fwrd'" oligonucleotides in the presence of Mn^{2+}, whereas the other three oligonucleotides did not show any cleavage (see Figure 32). A similar preference of Mn^{2+} instead of Mg^{2+} has been observed previously for $TraI_F$ from *E. coli* in cleavage experiments with ssDNA substrates (296). The size of the cleavage products was compared with a standard made of different polyT oligonucleotides of different lengths (data not shown). The cleavage products of the "GGI *oriT* 'fwrd'" oligonucleotide and the "F oriT 'fwrd'" oligonucleotide had lengths of 6-15 and 65-70 nucleotides respectively. A comparison of the putative cleavage sites did not show any sequence similarity between both oligonucleotides. Remarkably, cleavage of the "GGI *oriT* 'fwrd'" oligonucleotide was inhibited in the presence of ATP.

Further cleavage assays with several shorter oligonucleotides along the sequence of the "GGI *oriT* 'fwrd'" oligonucleotide and oligonucleotides with mutations in single nucleotides are required to determine the exact cleavage site and the specificity of TraI.

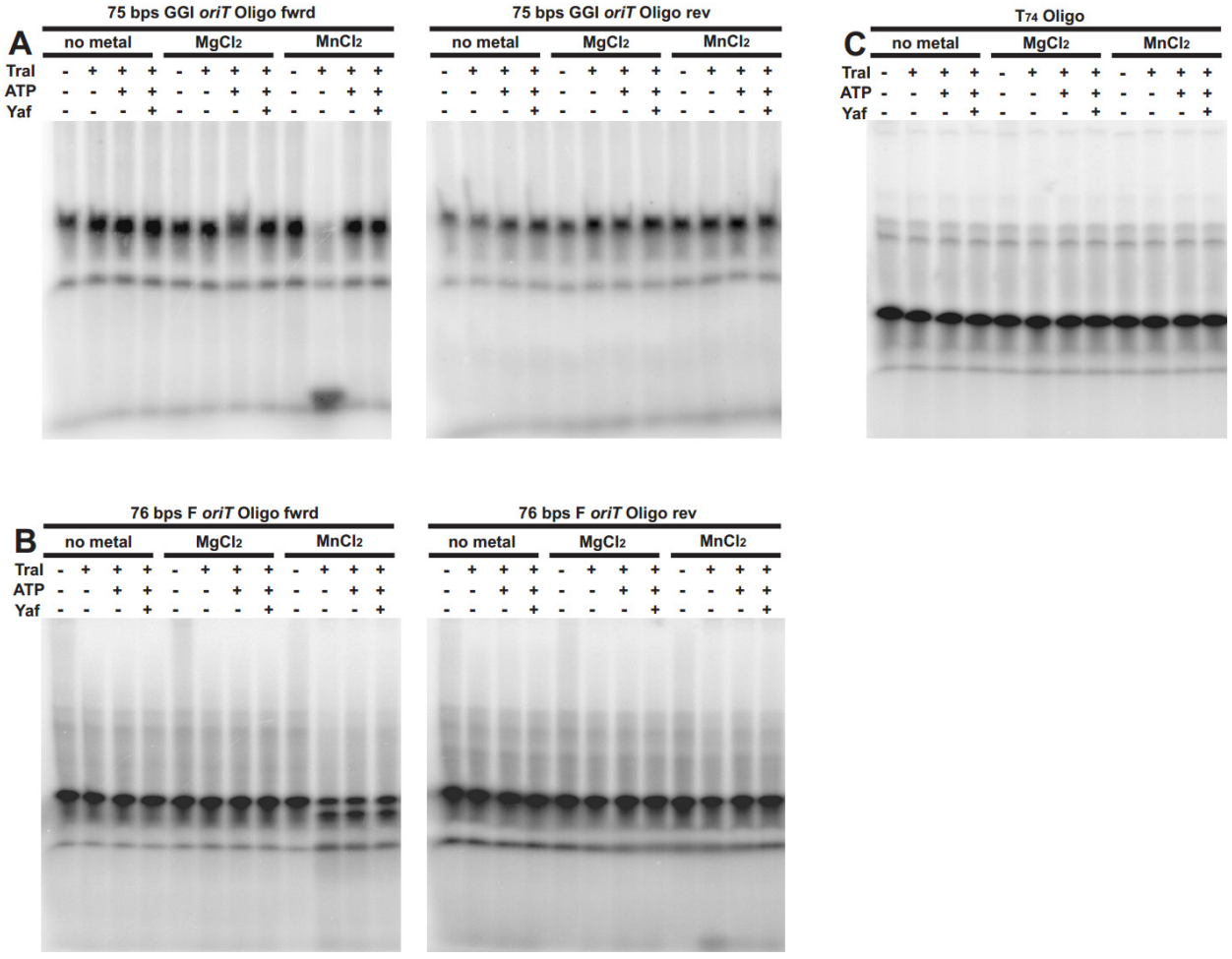


Figure 32: Autoradiographs of DNA cleavage assay with TraI and Yaf. The samples were analyzed with Urea PAGEs with 7.2% gels. The labeling of the different lanes represents the corresponding reaction conditions. The assays were performed with six different γ-P^{32}-labeled DNA substrates. (A) Cleavage reaction performed with complementary '75 bps GGI *oriT*' oligonucleotides; (B) Cleavage reaction performed with complementary '75 bps F-plasmid *oriT*' oligonucleotides; (C) Cleavage reaction performed with 'T_{74}' oligonucleotide (left) and a 555 bps PCR product containing the GGI *oriT* sequence (right)

4.5 Protein-protein interaction studies on the relaxase TraI, Yaf and the ATPase TraC

The studies on possible protein-protein interaction of TraI, Yaf, and TraC were performed in close collaboration with S. Rempel (a bachelor student of our group), whose bachelor project focused on the isolation and biochemical studies of the ATPase TraC of the T4SS from *N. gonorrhoeae*.

Different combinations of isolated proteins were co-incubated with magnetic His-beads and were eluted from the beads as described in Chapter 3.8.11. The different fractions from the elution of the magnetic beads were analyzed by SDS-PAGE. Putative protein-protein interaction will be indicated by co-elution of the different proteins, since only one of the two co-incubated proteins contains a His-tag. Possible protein-interactions of native TraI were tested with dimeric and tetrameric Yaf with an N-terminal His-tag as well as with N-terminal His-tagged TraC. A possible interaction of Yaf with the ATPase TraC was tested using native Yaf and N-terminal His-tagged TraC.

However, putative protein-protein interactions between TraI, Yaf, and/or TraC were not observed within the tested conditions. A similar experimental approach to study the interaction of TraI and Yaf was performed using Ni-NTA (Sigma-Aldrich) instead of magnetic beads for the binding of His-tagged proteins. Like observed for the magnetic bead material, also the obtained results from the Ni-NTA material do not indicate an interaction between TraI and Yaf.

Nevertheless, it has to be taken into account that a weak interaction between the different proteins could possibly be disturbed by various factors like e.g. buffering conditions or the treatment of the samples. For instance, TraI needs the presence of the detergent DDM to keep a soluble dimeric state in solution. The micelle formed by DDM could hinder possible protein interactions of TraI. Besides the possibly disturbing influence of DDM also the concentration of salts and ionic components could have a strong influence on protein-protein interactions and should be considered in the interpretation of the obtained results.

In addition to the performed pull down assays, a possible influence of TraI and Yaf on each other was tested in DNA cleavage assays and DNA binding experiments. The obtained results from the cleavage assays are in agreement with the results of the pull down assays. No indication of a putative influence of TraI and Yaf on each other could be observed. However, within the experimental set-up of the DNA cleavage assays it is not possible to detect a direct protein-protein interaction between TraI and Yaf.

5. Discussion

5.1 GGI-like T4SSs in Proteobacteria

GGI-like T4SSs show variations within different Proteobacteria

Generally, the GGI-like T4SSs that have been identified in several Proteobacteria within the current study show a highly conserved genetic organization of the different T4SSs components. However, a few exceptions show little variation in their genetic organization, which result from additional insertions. The largest insertions are present within the GGI-like T4SSs identified in γ-Proteobacteria, but also α-Proteobacteria show significant additional insertions. The most conserved GGI-like T4SSs can be found in β-Proteobacteria. The additional insertions in the identified T4SSs of α- and γ-Proteobacteria presumably derive from evolutionary adaptation processes within highly variable genomes.

The GGI of *N. gonorrhoeae* has been proposed to derive from an ICE, since it is surrounded by two *dif* sites that are recognized by the XerCD recombinase (198, 200). One of the two *dif* sites has been mutated to an imperfect *dif* site (189). The mutation within this *dif* site influences the recognition of the XerCD recombinase and leads to a strongly decreased excision and integration frequency of the GGI. The GGI-like T4SSs identified in other Proteobacteria are often surrounded by transposases or integrases, indicating a possible origin from a mobile genetic element similar to the GGI of *N. gonorrhoeae*. It is most likely that the GGI-like T4SSs derive from one ancestral T4SS which has been distributed via a mobile element or due to bacterial phages, since very often also many phage related proteins are in close proximity to the genes of the GGI-like T4SSs. Interestingly, only the conserved core genes necessary for DNA processing and the formation of the secretion channel are found in most of the identified GGI-like T4SSs. This indicates a clear tendency of these T4SSs to spread only essential core genes among different organisms. The bioinformatical analysis of the different GGI-like T4SSs does not give information about possible useful cargos within the spread of these systems, like e.g. antibiotic resistances or metabolic advantages.

Functional role of GGI-like T4SSs

Almost all identified species harboring a GGI-like T4SSs are pathogenic organisms, suggesting a possible function of the T4SSs in pathogenesis or in immune defense. A direct influence of the GGI encoded T4SS on the pathogenicity of *N. gonorrhoeae* has not been reported up to now, however, GGI versions encoding a specific *sac-4* allele have been associated with an increased risk of PID infection (113). Furthermore, it has been reported that the GGI encoded T4SS promotes a ton-independent intracellular survival of *N. gonorrhoeae* being an advantage for invasion and colonization of distinct tissues (196). However, it is unclear whether the GGI-like T4SSs secret additional virulence factors like it has been reported for many other T4SSs (102, 104, 121). No secreted proteins that possibly act as virulence factors have been identified in *N. gonorrhoeae* and it is currently assumed that the GGI encoded T4SS is a pure DNA secretion system. GGI-like T4SSs might be involved in surface modulation to evade the host immune system. For instance, *N. gonorrhoeae* strain MS11 has a mutated pilus subunit TraA and does not assemble T4SS pili. Other *N. gonorrhoeae* strains do not have this mutation and have the ability to form intact T4SS pili. However, the presence

of a pilus in the neisserial T4SS significantly influences the level of DNA secretion (PhD thesis E. Pachulec 2010). Besides a putative role within pathogenicity, it is very likely that GGI-like T4SSs are important for the spread of information via DNA secretion. The different species might benefit from the exchange of genetic information in acquiring antibiotic resistances or metabolic advantages that could easily be distributed via DNA secretion.

The obtained bioinformatical data does not give information about the functionality of the identified GGI-like T4SSs. Moreover, the large insertion in the GGI-like T4SSs found in α- and γ-Proteobacteria most likely lead to a loss of functionality, since operons with essential core genes are disrupted and could not be further coexpressed. The strict regulation of certain secretion system components within T4SSs is a crucial mechanism to enable a sufficient assembly of the secretion complex. The GGI-like T4SSs identified in β-Proteobacteria presumably are functional T4SSs since the essential core genes are not disrupted by additional insertions. The insertions between the *dtr* genes and the *mpf* genes most likely do not affect the functionality of the T4SSs, instead of that it is more likely that these insertions belong to a more ancestral T4SS and have been lost within *N. gonorrhoeae* and *N. meningitidis*. If this assumption will be true, it is most likely that the GGI-like T4SS ancestrally derives from one of the other β-Proteobacteria species and not from *Neisseria*, like it has been previously assumed. This implies that the T4SS of *N. bacilliformis* resembles a more ancestral T4SS. However, conclusions considering the functionality and the ancestral relationship of the GGI-like T4SSs have to be seen with caution as long as no further experimental and bioinformatical analyses have been performed.

Some GGI-like T4SS encode different classes of dtr genes and mpf genes

Some of the GGI-like T4SSs have a different composition of *dtr* genes and *mpf* genes. The T4SSs of *N. aromaticivorans*, the secretion system of *A. denitrificans* plasmid pALIDE201, and the T4SS of plasmid pAOV001 of *Acidovorx* encode *dtr* and *mpf* genes belonging to the MOB_F and the MPF_F family, whereas all other identified GGI-like T4SSs encode *dtr* genes belonging to the MOB_H family and *mpf* genes belonging to the MPF_F family. Such an exchange between different MOB and MPF families has also been reported for other T4SSs (111). Most of the identified mixed T4SSs are encoded on the chromosome and derive most likely from mobile genetic elements. These “original” secretion systems might have changed due to different evolutionary processes that finally resulted in a mix of different MOB and MPF families. Many of the mixed GGI-like T4SSs are surrounded by or even encode different phage related proteins. This gives rise to the assumption that the change of the *dtr* genes might result from phage integration events. Nevertheless, conclusions about the origin of the mixed GGI-like T4SSs have to be drawn with caution since several other factors might lead to an exchange of different MOB and MPF families.

5.2 Isolation of the coupling protein TraD

Soluble TraD could not be isolated using a recombinant overexpression approach in E. coli

The recombinant overproduction of the neisserial CP TraD led to the formation of large amounts of insoluble inclusion bodies, even though various overexpression conditions, different constructs and different *E. coli* host strains have been tested.

In general, T4SS CPs are integral inner membrane proteins with at least two TM domains. A few CPs have been reported to lack TM domains, but these proteins always occur together with small integral membrane proteins that mediate the attachment of the CPs to the cytoplasmic membrane (105). The CP of *N. gonorrhoeae* is, based on its conserved nucleotide binding motifs, classified into the group of $VirD4_{AT}$ like CPs (105). Taking the results from computational TM domain predictions of TraD into account, it is more likely that TraD clusters together with the class of TM domain-less CPs since no TM domain has been predicted for TraD. Instead, a hydrophobic patch could be identified within the N-terminal region of TraD, which most likely is responsible for the attachment of TraD to the cytoplasmic membrane. Considering these results, it might be possible that TraD needs an additional protein to enable an attachment or integration into the membrane. This assumption is strengthened by the close proximity of the *yaa* gene to the *traD* gene, which encodes a small predicted membrane protein with at least 4 TM domains. The possibility of a missing component that might be necessary for the successful attachment or integration of TraD into the cytoplasmic membrane could be one reason for the production of large amounts of insoluble TraD during overproduction. The membrane attachment or insertion of TraD could be triggered by an up to now unknown additional spatial positioning factor in *N. gonorrhoeae*, which is not present in recombinant overexpression host *E. coli*. This assumption is of special interest with regard to the low abundance of T4SS components, often with only one copy per cell. Furthermore, CPs oligomerize to homohexameric holoenzymes and resemble a similar shape like F_1F_0-ATPases (274). This implies the formation of several protein-protein interactions between CP subunits, indicating a possibility for an aggregation tendency of CPs while overproduction in high amounts. It is assumed that CPs assemble into their active homohexameric state after receiving a certain signal, which initiates the activation of substrate translocation (229). Like previously discussed in context with a putative missing spatial positioning component, also a missing oligomerization triggering signal might lead to the observed aggregation of TraD during the recombinant overproduction. In addition to oligomerization related challenges that need to be considered within the recombinant overproduction of TraD, also host specific peculiarities could result in the formation of inclusion bodies. For instance, the lipid composition of the cytoplasmic membrane could strongly influence the integration of a recombinant membrane protein into the host membrane, if reasonable differences in the lipid composition of both species exist. This might be possible for the recombinant overexpression of TraD from *N. gonorrhoeae* (β-Proteobacteria) in *E. coli* (γ-Proteobacteria), even though both belong to the group of Proteobacteria. Furthermore, the overexpression of *traD* results in the production of high amounts of protein, which could cause membrane insertion problems due to limited membrane space in the host organism leading to the formation of inclusion bodies. Besides this, the fast production of large protein amounts during the induction with IPTG can lead to a misfolding of the protein that ends up in the formation of inclusion bodies. The recombinant overproduction of membrane proteins is in general a challenging task and faces often problems of insolubility or misfolding of the produced

proteins. Also other T4SS CPs could not be isolated and purified as full-length proteins; only the soluble cytoplasmic part could be successfully overproduced and purified (e.g. $TrwB_{R388}$) (270, 274).

Denatured TraD could be purified from inclusion bodies but aggregates within refolding experiments

Denatured TraD was isolated and in part purified from inclusion bodies, but any further purification steps resulted in aggregation or loss of the protein. A few of the tested refolding conditions ended up with soluble protein, indicating a structural rearrangement of the denatured protein during refolding. However, it is not possible to draw conclusions about the correct folding state of TraD from the obtained data. A misfolding of the denatured protein during the refolding reaction most likely results in the aggregation of the protein, like it has been observed for TraD. Therefore, it is obvious to assume that TraD does not refold into its correct structural arrangement. The unpsecific binding of refolded and purified TraD to any kind of surface (e.g. different column material or reaction tubes) gives additional evidence for a misfolding of the protein. The strong surface attachment of TraD is very likely caused by surface-protein interactions of hydrophobic residues within TraD and charged surface material; possibly these interactions occur within the computationally predicted hydrophobic patch in the N-terminal domain of TraD. It is an overall observation that certain proteins, especially membrane proteins, tend to attach due to hydrophobic interactions to various surfaces (e.g. reaction tubes). Interestingly, even the presence of detergents did not prevent the surface attachment of TraD. Currently, there is no explanation for this contradictory behavior of TraD.

Membrane-bound TraD could not be solubilized from isolated E. coli membranes

Small amounts of membrane-bound TraD have been observed during numerous isolation attempts, but it was not possible to solubilize the membrane-bound TraD, even though several different detergents and solubilization conditions have been tested. Considering the observed data from different solubilization screens, it can be concluded that the membrane-bound TraD represents strongly aggregated protein. It is not clear from the obtained results, whether the membrane-bound TraD was integrated in small amounts into the cytoplasmic membrane during the overproduction, or if it results from persisting contaminations with inclusion bodies. Separation of the inner and outer membranes with sucrose density step centrifugation was used to remove possible inclusion body contaminations from the isolated membranes. Nevertheless, it cannot be absolutely excluded that small amounts of inclusion bodies were merged with the cytoplasmic membranes due to strong mechanical forces occurring during the cell breakage. The possibility of a mixed membrane contaminated with inclusion bodies explains the observed insolubility of the membrane-bound TraD.

5.3 Isolation and functional characterization of the hypothetical protein Yaf

Yaf forms different oligomeric states in solution

Yaf forms different oligomeric states in solution: the most dominant forms are dimers and tetramers, but also monomers and trimers have been detected. Interestingly, the presence of an N-terminal

His-tag strongly influences the equilibrium between dimeric and tetrameric Yaf and results in a switch of the oligomer ratio. The changes within the oligomerization pattern of Yaf in the presence of an N-terminal His-tag indicate that the N-terminal region of Yaf might be involved in protein-protein interactions. Considering a putative nicking accessory function of Yaf, the observed influence of the N-terminal region on protein-protein interactions is in contrast to reported data from other nicking accessory protein, which have C-terminal regions involved in protein-protein interactions and N-terminal regions facilitating the DNA binding (250, 259). However, from the obtained data it is not clear whether Yaf consists of two domains like other nicking accessory proteins, or if it resembles a one domain protein.

The functional role of the different oligomeric states of Yaf still needs to be elucidated further, to identify the final stoichiometry of the Yaf-holoenzyme. The current study revealed the dimeric and tetrameric forms of Yaf representing the active states of the protein, with a slightly stronger activity of the dimeric form. Many regulatory proteins with conserved RHH motifs, as well as nicking accessory proteins bind DNA in a dimeric form and tetramerize afterwards to bend the DNA either for the regulation of gene-expression or to enable access of the relaxase (215, 250, 260). The observed distribution of the DNA degradation activity within the different oligomeric states of Yaf is in agreement with reported data from RHH regulatory proteins and nicking accessory proteins (215, 259), even though no direct DNA binding of Yaf has been observed and the functional mechanism of Yaf still needs to be further studied.

Limited proteolysis revealed a high level of trypsin resistance of Yaf

Limited proteolysis experiments revealed a high grade of resistance (>20 µM) of Yaf against the proteolytic activity of trypsin. No significant differences have been observed between dimeric and tetrameric Yaf, even though a tendency of an increased resistance has been observed for tetrameric Yaf. The stability of Yaf against the proteolytic activity of trypsin is correlated with the structural exposure of accessible trypsin cleavage sites, since trypsin is an endoprotease that cleaves specific after the alkaline amino acids lysine and arginine. It is most likely that the oligomerization of Yaf protects the protein on a certain level against the proteolytic attack of trypsin. Besides the protection of Yaf due to its oligomeric assembly, also the structural arrangement of the monomers itself influences the accessibility of trypsin cleavage sites. The obtained data does not give reliable information about the domain organization of Yaf monomers, however, it is conclusive to assume that Yaf consist of only one domain facilitating the DNA degradation reaction and the oligimerization.

Limited proteolysis of different nicking accessory proteins revealed a common two domain organization of these proteins into an N-terminal DNA binding domain and a C-terminal protein-protein interaction domain (250, 259). Both domains have been reported to be highly resistance against the proteolytic activity of trypsin. It is obvious that trypsin starts to attack a protein initially on easy accessible sites that are often inter-domain regions forming exposed loop structures. Therefore, limited proteolysis is a commonly used method to characterize the domain organization of proteins. The reported data from the proteolytic digest of different nicking accessory proteins strengthens the assumption that Yaf resembles a one domain protein, since no second Yaf related fragment has been observed. This implies that Yaf combines both, the catalytic site for DNA degradation and the protein-protein interaction site within one domain, contrary to nicking accessory proteins.

Point mutations strongly influence the solubility and the overproduction yields of Yaf

The introduction of single point mutations into the sequence of Yaf strongly influences the solubility of recombinant overproduced Yaf and results in significantly reduced protein yields. The obtained data indicates a strong influence of the mutations on the folding and structural arrangement of Yaf. Considering the drastic changes on the solubility and the protein yields related to the introduction of point mutations, it is conclusive to assume a densely packed structure of Yaf that is very sensitive to changes in the amino acid assembly. It has to be considered that most of the introduced mutations were non conserved mutations to alanine and only a few of them were conserved mutations with similar structural amino acid properties (e.g. Y to F mutations). A possible influence due to changed biochemical and structural properties of the amino acids on the structural assembly of Yaf cannot be excluded, however, reduced protein yields and the formation of inclusion bodies have also been observed for conserved mutations (Y to F mutants).

Yaf binds to dsDNA substrates with low affinity

Evidences for a weak binding of dsDNA substrates have been observed within several DNA binding assays, but the obtained results have to be interpreted with caution. It has to be considered that maybe the tested *in vitro* conditions do not reflect the optimal DNA binding conditions of Yaf and further optimization of the assay might be necessary.

Yaf has a sequence unspecific metal dependent relaxation activity on supercoiled plasmid DNA

Isolated dimeric and tetrameric Yaf has a sequence unspecific relaxation activity on supercoiled plasmid DNA similar to TraI and *E. coli* Topoisomerase I. The relaxation activity of Yaf is dependent on divalent metal ions as cofactors and is strongly enhanced in the presence of Mn^{2+} and Co^{2+} in comparison to Mg^{2+}.

A possible contamination of the isolated protein can be excluded for the same reasons as described for TraI, namely the enhanced activity of Yaf and the reported reduced activity of Topoisomerase I in presence of Mn^{2+}. Additionally, the relaxation activity of Yaf has been observed for dimeric and tetrameric protein which has been purified via size exclusion chromatography, and a putative contamination with Topoisomerase I most likely would not be present in both fractions, since the size of Topoisomerase I is supposed to elute in a complete different volume (Topoisomerase I – 97.4 kDa; Yaf dimer – 34.2 kDa; Yaf tetramer – 70.4 kDa).

Currently, it is not known why the hypothetical protein Yaf has a DNA relaxation activity similar to that of the relaxase TraI. However, it might be possible that Yaf is involved in the DNA processing mechanism not as an NAP like previously assumed, but as a protein that supports the activity of TraI, which seems to lack a distinct helicase domain. Considering the relaxation activity and the DNA degradation activity of Yaf, it is obvious that Yaf has the ability to cleave DNA; therefore, Yaf might also be involved in the termination reaction during DNA processing. Nevertheless, it is more likely that the termination reaction during the DNA processing reaction in *N. gonorrhoeae* is mediated by the relaxase TraI, like it has been reported for other T4SSs (30, 239, 240).

Several lines of evidence give rise to the speculation that Yaf possibly has a regulatory function in *N. gonorrhoeae*, even though the detailed mechanism is up to now not known. For instance, Yaf might act in a control mechanism for the correct excision and circularization of the GGI from the chromosome, similar to RDFs (Recombination Directionality Factors) within ICEs (179, 180). The neisserial GGI has been reported to excise in a low frequency (due to an imperfect *dif* site) from the chromosome and forms a circular intermediate (189). The imperfect *dif* site possibly leads to a certain amount of incorrect excised GGIs and the cells would need a mechanism to reintegrate or even degrade the incorrect excised DNA, which might be enabled by the observed nuclease activity of Yaf. Evidence for this assumption is given with the specific nuclease activity of Yaf that is restricted to linearized DNA fragments. Furthermore, it has been reported that some RDFs show an *E. coli* Topoisomerase I like DNA relaxation activity *in vitro*, like it has been observed for Yaf (180).

Yaf degrades linear ssDNA and dsDNA fragments in a sequence unspecific metal dependent reaction

Dimeric and tetrameric Yaf degrade ssDNA and dsDNA fragments in a sequence unspecific and metal dependent reaction. The DNA degradation activity of Yaf is dependent on the metal cofactors Mn^{2+} and Co^{2+}, and is not observed in the presence of Mg^{2+}. Interestingly, the DNA degradation activity of Yaf is restricted to linearized DNA fragments and circular plasmid DNA is not degraded. This might indicate a putative exonuclease activity of Yaf, since there is evidence that Yaf initiates the degradation from the edges of the DNA fragments, which would also explain why Yaf does not degrade plasmid DNA. However, further experimental studies are necessary to prove this assumption.

A possible background activity resulting from a contamination with an exonuclease of *E. coli* can be most likely excluded as, on the one hand, the DNA degradation activity of Yaf can be observed for ssDNA and dsDNA and is dependent on Mn^{2+}, whereas the exonucleases of *E. coli* are mostly either specific to ssDNA, to dsDNA or to RNA, or they require Mg^{2+} or Ca^{2+} as metal cofactors. Furthermore, the DNA degradation activity of Yaf can be observed for two different oligomeric forms of Yaf that elute within size exclusion chromatography in volumes representing complete different molecular masses and it is very unlikely that a possible contamination of an exonuclease will elute in exact the same volumes as dimeric and tetrameric Yaf.

The functional role of the nuclease activity of Yaf still needs to be elucidated in further studies. However, considering the assumption that Yaf might have a regulatory function in *N. gonorrhoeae*, two different mechanisms might be possible. One might function as a control mechanism for the correct excision of the GGI, like it has already been discussed in detail in the previous passage. Another possible role of Yaf might be the regulation of the DNA secretion level under oxidative stress conditions. Evidence for this assumption is given in the observed decreased DNA cleavage activity of TraI in presence of Yaf and the Mn^{2+} dependent DNA degradation activity of Yaf. Such a regulative mechanism would facilitate access to an important phosphate source, which is given in the ssT-DNA that could be used for the production of either pyrophosphate (PP_i) or poly-phosphate (polyP). Both, PP_i and polyP are known to act together with Mn^{2+} as antioxidants (297, 298). If *N. gonorrhoeae* secretes DNA in a constitutive process like it is currently assumed, such a mechanism might be of special interest for the detoxification of reactive oxygen species (ROS) during oxidative stress.

5.4 Isolation and functional characterization of the relaxase TraI

The MOB_H relaxase TraI resembles certain features of membrane proteins

The isolation and purification of recombinant overproduced TraI showed a strong tendency of TraI to aggregate during the purification process. Various overproduction conditions, different expression constructs and purification procedures have been tested to optimize the isolation of TraI, but finally ended up with insoluble aggregated protein. Recently, it was for the first time possible to isolate and purify soluble TraI in presence of the detergent DDM.

Most likely, the aggregation tendency of TraI rises from several factors that together have an additive effect resulting in the aggregation of the protein. For instance, the enzymatic activity of TraI comes along with a high DNA binding affinity that could cause an aggregation of TraI already in the initial isolation step during cell breakage, since the protein strongly tends to bind sheered DNA fragments resulting from the high forces occurring during cell breakage. To avoid an aggregation due to multiple protein binding on sheered DNA fragments, DNAseI is added before the cells were broken to degrade the DNA immediately after release. However, DNAseI treatment alone was not sufficient to avoid the aggregation of TraI, indicating either a protection of the DNA against the degrading activity of DNAseI by bound TraI, or other factors are responsible for the observed aggregation. Taking into account that TraI is a relatively large protein consisting of 850 amino acids (94.7 kDa) that is most likely secreted via the T4SS, like it has been reported for several other relaxases, it is obvious that TraI needs a certain structural flexibility to enable a partial unfolding which is necessary for the secretion through the secretion complex (103). This property of TraI might result in an aggregation of the protein during the recombinant overproduction in *E. coli*, since the protein could easily be misfolded within the fast folding procedure. However, the most probable reason for the observed aggregation of TraI is the hydrophobic N-terminus. The accumulation of hydrophobic residues in the N-terminal region of TraI tends to aggregate with other N-termini in a non-polar solution. This aggregation can be circumvented by the use of detergents that keep the hydrophobic N-termini separate from each other by forming micelles around TraI, like it is a standard procedure for membrane proteins. The fact that TraI stays soluble during the purification in presence of the detergent DDM strengthens the assumption that the hydrophobic N-terminus of TraI is causing the observed aggregation. Considering the purification properties of TraI and the fact, that reasonable amounts of TraI can be detected in isolated membranes, TraI shows *in vitro* some features that are typically found in membrane proteins.

TraI forms stable dimers in solution

Experimental data reveals the formation of stable TraI dimers in solution. In size exclusion chromatography TraI eluted in a volume corresponding a molecular mass of 278 kDa, which represents the approximate mass of a TraI dimer surrounded by a DDM micelle (189.4 kDa + 70 kDa). Data obtained from Blue Native PAGE indicate a molecular mass of TraI of 260 kDa, which corresponds to the mass of a TraI dimer and is in accordance with the data obtained with size exclusion chromatography.

The observed results are in agreement with reported data from several other relaxases forming dimers in solution (239, 240, 299). Most relaxases stay covalently bound to the single stranded transfer DNA after the cleavage reaction and a second relaxase protein is needed to terminate the DNA processing reaction (212, 214). However, it is still under discussion if this hypothesis can be generalized for all relaxase families since many relaxases differ in their catalytic site (one or two catalytic tyrosines) and examples have been reported that do not bind covalently to the ssT-DNA (236). Experimental data obtained from DNA binding and competition assays indicate that isolated TraI from *N. gonorrhoeae* does not bind DNA in a covalent reaction *in vitro*, however, it cannot be excluded that TraI possibly attaches covalently to the DNA after the initial cleavage reaction. Since no biochemical data is available for the MOB_H prototype relaxases TraI from *N. gonorrhoeae*, it is not known, whether the DNA processing reaction is terminated with a second relaxase protein, or if one relaxase protein is sufficient to enable initial cleavage and termination. The experimental data obtained within the current study indicate the formation of stable TraI dimers, which most likely represent the active form of TraI. These data suggest the involvement of a second TraI protein in the termination of the DNA processing reaction. Sequence analysis and mutagenesis studies coupled with DNA secretion assays revealed at least three conserved tyrosines in TraI that might form the catalytic site, but up to now the active tyrosine of TraI has not been identified (202). Therefore, it is not known whether one or two catalytic tyrosines are involved in initializing and terminating the DNA processing reaction in TraI. Considering the observed dimer formation of TraI in solution, it is most likely that one TraI protein initiates the cleavage reaction and a second one terminates the reaction, however, this assumption need to be proven experimentally.

TraI can be separated into three different domains

Limited proteolysis with trypsin and ProteinaseK revealed at least three distinct domains in TraI. In lower trypsin concentrations the formation of two initial domains has been observed, representing molecular masses of 80-85 kDa and 35 kDa on SDS-PAGE gels. In high Trypsin concentrations, a third domain with a molecular mass of 55 kDa occurred on the gel and the 80-85 kDa starts to disappear.

Generally, relaxases consist of at least two different domains, a relaxase domain for the DNA cleavage reaction and a helicase domain for the unwinding of the ssT-DNA (212). TraI from *N. gonorrhoeae* contains besides the N- terminal relaxase domain an additional C-terminal DUF1528 domain that often can be found in proteins with an N-terminal metal-dependent phosphohydrolase (HD) region (e.g. helicases or relaxases). In average, DUF1528 domains consist of 100 amino acids with an approximate molecular mass of 10-12 kDa. The distinct function of this domain is still unknown, but it is assumed that the DUF1528 domain of the neisserial relaxase TraI is involved in protein-protein interactions (202). The DUF1528 domain of TraI has a predicted molecular mass of 11.8 kDa.

The observed formation of at least three different domains in TraI is consistent with reported data from other relaxases, even though no data is available on MOB_H relaxases (211, 300). The large domain with a molecular mass of 80-85 kDa represents most likely the N-terminal relaxase domain of TraI. This assumption is based on bioinformatical data that indicate a large N-terminal TraI_2 relaxase domain with a length of 328 amino acids (42.5 kDa). The relaxase domain itself does not completely represent the observed 80-85 kDa domain, indicating the presence of an additional middle part within TraI. Bioinfomatical alignment data (not shown) revealed the presence of an in part conserved

middle domain within many TraI_2 relaxases. This conserved middle domain most likely represents the remaining molecular mass of the obtained fragment. The occurrence of an additional 55 kDa fragment at higher trypsin concentrations accompanied with the disappearance of the large 80-85 kDa fragment, indicating a separation of the relaxase domain and the remaining conserved middle part. However, the obtained experimental results are inconclusive about whether the 55 kDa fragment is the N-terminal relaxase domain or if this fragment represents the conserved middle part of TraI. Further experiments (e.g. peptide finger print analysis or N-terminal sequencing of the obtained domains) need to be done to gain information about the sequence composition of the different domains, to identify their specific location within the sequence of TraI. The domain with a molecular mass of 35-17.5 kDa most likely represents the C-terminal DUF1528 domain, which has a predicted length of 105 amino acids (11.8 kDa). Considering the reported domain organization of TraI with an N-terminal relaxase domain and a C-terminal DUF1528 domain and taking into account that the proteases attacks at first easy accessible sites, e.g. loop regions between two domains, it is most likely that the observed 35-17.5 kDa fragment represents the C-terminal DUF1528 domain since this fragment occurs already in low trypsin concentrations. The difference within the predicted mass of the DUF1528 domain (11.8 kDa) and the obtained mass of 35-17.5 kDa most likely results from the cleavage position of the protease, which could be in a significant distance to the actual domain since the attack occurs at an easily accessible structural region between two domains.

Identification of catalytic residues involved in the DNA cleavage reaction of TraI

The relaxase domain of the MOB_H relaxase TraI contains several conserved sequence motifs typically found in other relaxase families, but also includes an altered histidine motif and conserved HD domains that have not yet been identified in other relaxase families (110, 202).

The structural modeling program Phyre2 was used to obtain a model of the possible structure of TraI by comparing the amino acid sequence of TraI with available structures with similar sequences. A ribbon model of the modeled TraI structure is given in Figure 33. The overall modeled structure of the neisserial TraI has a similar shape like the crystal structures of $TraI_F$, $TrwC_{R388}$, and $MobA_{RSF1010}$ (227, 245, 246). The modeled structure of TraI has a handcuff like shape and resembles the typical structural arrangement of at least five α-helices and an anti-parallel two stranded β-sheet forming the binding cleft with the catalytic site.

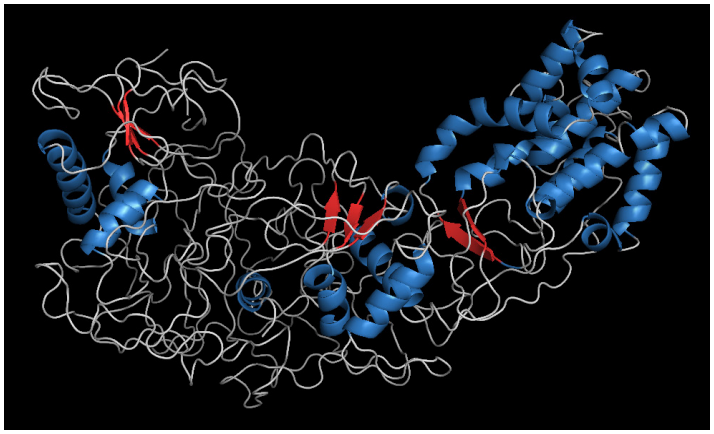


Figure 33: Side view on the structural ribbon model of TraI from *N. gonorrhoaea*. The structure has been modeled using the Phyre2 software. α-helices are colored in blue, β-sheets are colored in red and coiled structures are represented as grey line. The highest conservation has been identified within the N-terminal relaxase domain that resembles the typical handcuff like structure at the right part of the model.

Most likely, the active site of TraI is located within the binding cleft and the DNA binding is mediated via a similar "knob-into-hole" mechanism like it has been described for $TraI_F$, $TrwC_{R388}$ and $MobA_{RSF1010}$ (227, 245, 246). The altered histidine motif ($H^{106-108}$), as well as the conserved D^{120}, presumably represent the metal binding site and coordinate the metal ion in close proximity to the catalytic tyrosine. A spherical space filling model of TraI with color-labeled conserved residues is shown in a top view in Figure 34. The H-triad is indicated as yellows spheres and locates almost centrally within the cleft-vale. The conserved D^{120} is labeled in blue color and locates within an α-helix close to the histidine triad. The conserved Y^{93}, which has been reported to be essential for DNA secretion, is labeled in green and seems to be in close proximity to the histidine triad and D^{120}. However, Y^{93} seems to be buried within the structure and represents probably not the catalytic residue, since it seems not to be surface exposed. The second tyrosine (Y^{201}) that has been reported to have significant influence on DNA secretion is labeled in orange and is located on the edge of the opposite site of the binding cleft, in far distance to the metal ion coordination site. Therefore, Y^{201} probably does not represent the catalytic tyrosine as well. The third conserved tyrosine (Y^{212}; labeled in red color) does not have any significant influence on DNA secretion, but represents most likely the catalytic site, since it localizes surface exposed in close proximity to the metal ion coordination site.

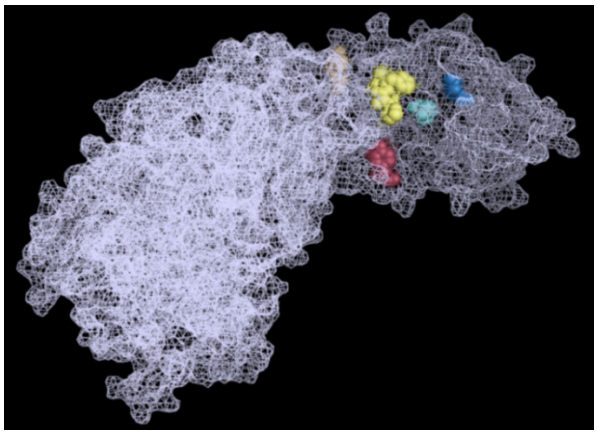


Figure 34: Top view on the spherical space filling model of the modeled structure of TraI from *N. gonorrhoeae*. Important catalytic residues putatively forming the active site are color labeled. The conserved histidine motife ($H^{106-108}$) are indicated as yellow spheres, the conserved D^{120} is colored in blue. The three conserved tyrosines are colored in green (Y^{93}), orange (Y^{201}) and red (Y^{212}).

A surface model of the modeled TraI structure gives additional information about the surface accessibility of the conserved residues and is depicted in Figure 35. The color labeling of the conserved residues corresponds to the previous description. The left model represents a site view with the metal ion coordination site and the conserved Y^{212} and illustrates the structural arrangement of the conserved putative catalytic residues within the binding cleft. On the opposite site of the binding cleft locates a small surface exposed part of Y^{201} (orange color) in far distances to the metal ion coordination site. The model on the right site represents a top view onto the binding cleft and indicates that only Y^{212} seem to be surface exposed, and the green labeled Y^{93} buried in the structure. However, Y^{212} probably needs to perform conformational changes to enable a closer localization next to the metal ion coordination sites, which might represent the active state of TraI.

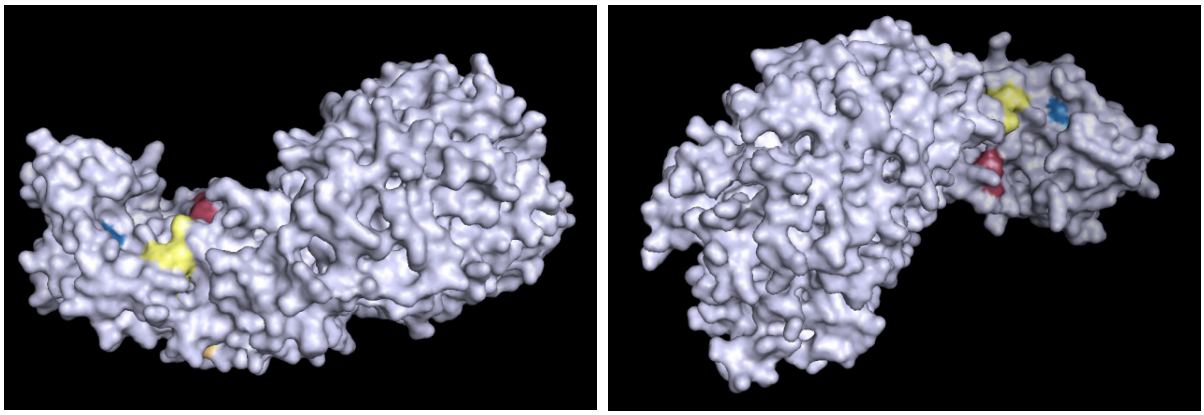


Figure 35: Spherical surface model of the modeled structure of TraI from *N. gonorrhoeae*. Important active site residues are color labeled. The conserved histidine motife ($H^{106-108}$) are indicated as yellow spheres, the conserved D^{120} is colored in blue. The three conserved tyrosines are colored in green (Y^{93}), orange (Y^{201}) and red (Y^{212}). Left: side view; Right: top view

Three conserved tyrosines have been identified in the relaxase domain of TraI (Y^{93}, Y^{201}, and Y^{212}), with one of them resembling the active residue for catalyzing the cleavage reaction of TraI (202). Mutagenesis screens and DNA secretion assays revealed that Y^{93} is important for DNA secretion and mutation to phenylalanine result in an abolished DNA secretion (202). A phenylalanine mutant of Y^{201} also shows significant effects on the DNA secretion activity of the T4SS of *N. gonorrhoeae* and results in a reduced secretion level (half of the standard level) (202). The mutation of the third conserved Y^{212} did not influence the levels of DNA secretion and seems to have no significant effect on the function of TraI (202). Besides the conserved tyrosines, also the altered histidine motif (H-triade) was mutagenized and has been tested for DNA secretion, but both, H^{106} and H^{108} mutants do not influence the DNA secretion levels significantly (202). The conserved HD domains have been assumed to be involved in the coordination of the metal ion, but until now no experimental evidence on this has been reported for TraI (202). However, mutations within H^{161} and D^{162} do not affect DNA secretion levels, whereas, a mutation of D^{120} leads to significant reduced secretion levels that are comparable to the results described for the Y^{201} mutant (202).

The experimental approach of measuring the DNA secretion activity the T4SS of *N. gonorrhoeae* does not give reliable information about the catalytic site within TraI and has to be interpreted with caution since *N. gonorrhoeae* tend to autolyze and high background signals have to be considered during the experiment. Furthermore, the mutations can cause structural rearrangements or misfolding of TraI resulting in an inactive protein, even though the substituted amino acids have biochemical and structural similarity. Therefore, the results of the described DNA secretion assays can be seen as an indication for a possible active site within TraI, and further experiments like e.g. DNA cleavage assays need to be done to gain detailed information about the active residues involved in the DNA processing reaction.

DNA binding properties of TraI

The current study has shown that TraI binds in a non-covalent sequence unspecific reaction to dsDNA and ssDNA fragments, with a preference to ssDNA. A minimal binding frame of 9 nucleotides has been identified for ssDNA fragments.

The minimal binding frame of at least 9 nucleotides most likely represents the length of the DNA strand that is needed to enable sufficient binding within the DNA binding site in the cleft. Other relaxases have been reported to bind DNA with a "knob-into-hole" mechanism using binding pockets distributed across the distance of the binding cleft, indicating the necessity of a certain minimal fragment length for DNA binding (232). Without having a structural prove it is very difficult to speculate about the DNA binding site of TraI, however, it is most likely that TraI binds DNA in a similar mechanism like other relaxases. Interestingly, TraI binds DNA in a compatible reaction demonstrating the non-covalent DNA binding activity. The obtained data does not give information about whether TraI binds generally in a non-covalent mechanism to DNA, or if the cleaved transfer DNA is bound in a covalent reaction and only the initial binding to the DNA fragments is a non-covalent reaction. It has to be taken into account that other relaxases have been reported to bind cleaved DNA in a non-covalent reaction, similar to the obtained data of TraI (236). If the neisserial TraI indeed binds also the processed transfer DNA in a non-covalent reaction, it is most likely that the termination reaction of TraI is mediated probably by the same tyrosine that is already involved in the initial DNA cleavage reaction. Nevertheless, further experiments to characterize the DNA binding of TraI within the cleavage reaction need to be done to elucidate the catalytic residues and the mechanism of DNA processing of the MOB_H relaxase TraI. The sequence unspecific DNA binding of TraI in the initial DNA binding reaction has also been reported for other relaxases and is in compliance with common literature (301). Very often other proteins, which probably have not yet been identified within *N. gonorrhoeae*, are necessary to enable a sequence specific DNA cleavage reaction of relaxases (210, 218). The preference of TraI to bind ssDNA fragments is in agreement with the commonly believed model that relaxases stay bound to the cleaved ssT-DNA and triggers the nucleo-protein complex to the secretion complex. Very often relaxases are also secreted together with the ssT-DNA into the recipient organism (103). The presence of 5' protected ssDNA in the culture medium has been reported for *N. gonorrhoeae*, indicating the possibility for a secretion of TraI possibly bound to ssDNA (202). Considering the enzymatic function of relaxases with special regard to their DNA binding properties and the fact that they stay bound to the ssDNA while secretion to the recipient cell, it is obvious that ssDNA substrates are preferably bound by these enzymes.

TraI relaxases supercoiled plasmid DNA and cleaves ssDNA fragments in a metal dependent and sequence unspecific reaction

A sequence unspecific relaxation activity on supercoiled plasmid DNA containing either the predicted *oriT* region of the GGI or the *oriT* region of the F-plasmid from *E. coli* has been observed for TraI. The relaxation activity of TraI is a metal dependent reaction using Mn^{2+} and Co^{2+} as preferred metal cofactors, and is very similar to the enzymatic activity of *E. coli* Topoisomerase I.

Considering the recombinant overproduction of the neisserial TraI in *E. coli*, a possible contamination of the isolated TraI with *E. coli* Topoisomerase I cannot be excluded. However, the observed

relaxation activity of TraI differs in several parts significantly from the activity of *E. coli* Topoisomerase I. The main differences between both enzymes are the preferred metal cofactors and the formation of characteristic reaction intermediates. Topoisomerase I has been reported to have a significant reduced activity in the presence of Mn^{2+} and Co^{2+} compared to the activity in the presence of Mg^{2+}, whereas the activity of TraI is strongly increased in the presence of Mn^{2+} and Co^{2+} and is decreased in the presence of Mg^{2+} (293). In addition to the preferred metal cofactor, the reaction of *E. coli* Topoisomerase I and TraI differs significantly in the formation of reaction intermediates. TraI does not form characteristic intermediates, whereas *E. coli* Topoisomerase I form a typical and commonly reported ladder-like intermediate pattern within agarose gel electrophoresis (302). Taking the observed differences of *E. coli* Topoisomerase I and TraI into account, it is clearly demonstrated that the observed relaxation activity of isolated TraI *in vitro* descended not from a contamination with Topoisomerase I, but represents a characteristic enzymatic function of TraI. Actually, the observed relaxation of supercoiled plasmid DNA is a typical property of relaxases and has been reported for many of them (299, 303). Relaxases generally cleave DNA (and relax supercoiled DNA) in a sequence specific reaction often supported by other relaxosome components that mediate the sequence specificity, however, TraI relaxases supercoiled plasmid DNA sequence unspecific. Most likely, the specificity of the relaxation reaction of TraI is triggered by an up to now unknown component, like e.g. an uncharacterized nicking accessory protein. Even though no relaxosome formation has been reported for TraI up to now, the formation of a putative relaxosome cannot be excluded. Possible evidences for the formation of a putative relaxosome are several small uncharacterized proteins that encode within the same operon like *traI* and *traD*. A very promising candidate for a putative nicking accessory protein is the small hypothetical protein Yaf, whose function is initially characterized within the current study. Another possible reason for the sequence unspecific relaxation activity of TraI might be a missing important part within the predicted *oriT* region of the GGI in the used constructs, since often *oriT* regions are very large sequence parts including several recognition sites for nicking accessory proteins, spatial positioning proteins, and the relaxase that together enable a sufficient sequence specificity within T4SSs. Finally, it has to be considered that the DNA processing mechanism of TraI has not been characterized up to now and no biochemical data of TraI is available. It is most likely that the sequence specific relaxation activity of TraI *in vitro* is caused by a missing additional component, which is presumably present in the host *in vivo* (e.g. TraD), or the predicted *oriT* region is too short and is possibly wrong annotated.

TraI has, in addition to the previously described relaxation activity, a metal dependent and sequence unspecific cleavage activity on ssDNA substrates. However, the observed cleavage activity is restricted to ssDNA fragments with secondary structures and has not been observed with poly-T ssDNA substrates. The cleavage activity of TraI is, similar to its relaxation activity, dependent on the two metal cofactors Mn^{2+} or Co^{2+}. The cleavage activity seems to be slightly reduced in presence of ATP and Yaf, indicating a possible regulatory function of Yaf, which will be discussed later in detail. It has been observed that TraI cleaves ssDNA substrates containing either the *oriT* region of the F-plasmid from *E. coli* or the *oriT* region of the GGI in a strand specific reaction, which is in agreement with common observations from other characterized relaxases, even though TraI differs from other relaxases in its sequence unspecific cleavage reaction (227, 304). The most probable reason for the sequence unspecific cleavage reaction is either a missing nicking accessory protein or the lack of certain recognition sites in the GGI *oriT* region containing oligonucleotide. However, this is the first report of a DNA cleavage activity of the uncharacterized MOB_H prototype relaxase TraI and it has to be taken into account that no comparable data of this novel relaxase family is available. To gain

further information about the DNA cleavage reaction of TraI it is necessary to optimize the cleavage assay and additional ssDNA substrates, as well as dsDNA substrates with different *oriT* regions need to be tested. Interestingly, TraI has been observed to bind to poly-T ssDNA substrates, but within the currently used cleavage assay no cleavage of these ssDNA substrates has been observed. This indicates a dependency of the cleavage activity of TraI on either secondary structures in the DNA substrates or a certain base composition. Many relaxases often nick the DNA at a distinct scissile phosphate that is mostly located between a "GC" within the *nic* site (212). It is very likely that the lack of a scissile phosphate exposed via the "GC" within the *nic* site is the reason for the not cleaved poly-T substrate. Further experiments are necessary to identify the minimal *nic* site components for the cleavage reaction of TraI and to identify whether secondary structures or the sequence itself are necessary to enable cleavage. However, the obtained data from the DNA cleavage assay excludes a possible role of Yaf as nicking accessory protein and give rise to the assumption of a putative regulatory function.

5.5 TraI and Yaf favor the extraordinary metal cofactor Mn^{2+}

Both, TraI and Yaf have been observed to favor Mn^{2+} as metal cofactor within DNA cleavage, DNA relaxation, and DNA degradation assays. The enzymatic activity of TraI is strongly enhanced with Mn^{2+} as metal cofactor compared to the activity level observed with Mg^{2+}. Furthermore, the DNA cleavage activity of TraI only occurs in presence of Mn^{2+} and could not be observed with Mg^{2+}. The relaxation activity of Yaf is, like observed for TraI, enhanced in presence of Mn^{2+}, but could be also observed with Mg^{2+}. The nuclease activity of Yaf is, however, strongly dependent on Mn^{2+} and has not been observed with Mg^{2+}. Both proteins show a similar enhanced enzymatic activity with Co^{2+}, which most likely is not the preferred natural metal cofactor, since the intracellular concentration of Co^{2+} in bacteria (5-10 µM) is much lower than the Co^{2+} concentration used in the *in vitro* assays (5 mM) (305).

The concentration of Mn^{2+} in different human tissues differs significantly within a range of nM up to µM and it has been commonly accepted that Mn^{2+} is an important triggering signal and virulence factor for many pathogenic organisms (306-308). Mn^{2+} significantly influences the regulation of several metabolic enzymes, as well as certain virulence factors (e.g. surface structures involved in adhesion) and plays an important role in the detoxification of ROS and RNS (reactive nitrogen species). During infection and colonization *N. gonorrhoeae* prefers to accumulate in human tissues with high Mn^{2+} concentrations, and proteomic studies of *N. gonorrhoeae* strain 1291 under different Mn^{2+} stress conditions revealed the important functional role of Mn^{2+} in the detoxification mechanism of ROS in *N. gonorrhoeae*. It has been reported that an accumulation of Mn^{2+} in *N. gonorrhoeae* results in an increased resistance to oxidative stress and reduces cell killing via post-translational modifications (306). The Mn^{2+} dependent detoxification of ROS is mediated either by Mn^{2+}-PP_i or Mn^{2+}-polyP complexes that act as non-enzymatic antioxidants (297).

The intracellular Mn^{2+} concentration within different bacterial species has been reported with µM up to mM levels (32 mM in *Lactobacillus plantarum*) (307, 309). The Mn^{2+} concentration used within *in vitro* activity assays of TraI and Yaf are in the possible physiological range and represent almost natural conditions. The reason why TraI and Yaf prefer Mn^{2+} as metal cofactor is up to now not

known and need to be further studied. Most likely, the concentration of Mn^{2+} in *N. gonorrhoeae* is within such a range that it is more efficient to use Mn^{2+} instead of Mg^{2+}, which is mostly used as metal cofactor in other T4SS relaxases (306, 310). Both metals have very similar biochemical properties and their most significant difference is the ion size (Mg^{2+} has an atomic radius of 1.72 Å, Mn^{2+} has an atomic radius of 1.79 Å) (www.environmentalchemistry.com). Therefore, it is easy to exchange both metals as enzymatic cofactors and the MOB_H relaxase TraI uses in the following Mn^{2+} instead of Mg^{2+} as metal cofactor.

Considering the fact that Mn^{2+} strongly influences the expression and regulation of several cellular components in *N. gonorrhoeae* and the obtained Mn^{2+} dependency of Yaf and TraI, it could be suspected that the T4SS of *N. gonorrhoeae* is possibly regulative influenced by Mn^{2+}, possibly in context with a detoxification mechanism during oxidative stress. To prove this assumption, further studies under oxidative and manganese stress conditions have to be done and a release of single phosphate resulting from the nuclease activity of Yaf should be tested.

5.6 TraI, Yaf, and TraC do not interaction in pull down assays *in vitro*

Possible protein-protein interactions between the relaxase TraI, the hypothetical protein Yaf and the ATPase TraC were tested with pull-down assays under different conditions. However, no interactions between Yaf and TraI, Yaf and TraC, as well as TraI and TraC have been observed. Possible interactions between TraI and Yaf were tested additionally within DNA cleavage assays, but no direct interaction between both proteins has been observed, even though the cleavage activity of TraI seems to be reduced in the presence of Yaf. Assuming a putative functional role of Yaf (e.g. as nicking accessory protein or as spatial positioning factor) within the DNA processing reaction of the neisserial T4SS, it is most likely that Yaf interacts with the relaxase TraI and possibly also with the core complex ATPase TraC. Several examples for protein-protein interactions within relaxosome components and the secretion complex have been reported (222, 255). For instance, it has been reported that the spatial positioning protein $TraM_F$ strongly interact with the coupling protein $TraD_F$, and the nicking accessory protein complex $VirC1_{AT}$/$VirC2_{AT}$ interacts with the relaxase complex $VirD1_{AT}$/$VirD2_{AT}$ (220, 255). It is therefore to be expected that Yaf interact most likely at least with TraI, if it has a functional role as putative nicking accessory protein. If Yaf would fulfill a regulatory function in *N. gonorrhoeae* it might be possible that is does not directly interact with TraI and TraC. Considering the obtained data, it is more conclusive to assume a putative regulatory role of Yaf, since there is evidence for a decrease of the cleavage activity of TraI in presence of Yaf and no direct protein-protein interaction between both proteins has been observed.

The possibility of having suboptimal *in vitro* reaction conditions has to be taken into account and maybe the experimental set up has to be changed to obtain possible protein-protein interactions. Other promising approaches to analyze putative interactions would be size exclusion chromatography or Surface Plasmon Resonance measurements. These assays would also allow the use of native proteins, whereas the pull down assay needs at least one protein with a tag. Considering the strong influence of the N-terminal His-tag on the oligomerization of Yaf, it is most likely that the N-terminal domain is involved in protein-protein interactions and possibly the His-tag

abolishes an interaction with TraI or TraC. Using another experimental approach with native proteins would exclude the possibility of a disturbing effect of a tag.

6. Future perspectives

The coupling protein TraD

Unfortunately, it was not possible within the current study to isolate soluble TraD for a biochemical characterization of this protein. Therefore, the main focus of future studies on TraD should be the production of soluble protein. One approach could be the overproduction in a different host organism, e.g. the *Bacillus* species, *Lactococcus lactis* or *Pichia pastoris*. Another approach for the isolation of soluble protein could be the fusion of TraD with small stabilizing proteins at its N- and C-terminus. This could be e.g. fusions with different combinations of the two small proteins YaiN and YbeL, which have been reported to optimize the integration and overproduction of membrane protein in recombinant *E. coli* expression systems (311). Also the fusion with TrxA (thioredoxin) could possibly increase the solubility of TraD and has been reported to circumvent the formation of inclusion bodies while overproduction in *E. coli* (312). Finally, a third possibility could be a fusion with GFP, which has been reported to optimize the membrane integration of proteins in *E. coli*. An advantage of the GFP-fusion approach would be the possibility of monitoring the correct insertion into the membrane by detecting the GFP signal (313).

To obtain soluble TraD it could be also necessary to coexpress *traD* together with other genes from the *dtr* opreon, most likely *yaa*, which encodes a small polytopic membrane protein and is located directly next to *traD*. This approach is based on the fact that several coupling proteins need additional small polytopic membrane proteins to integrate and attach to the membrane, especially coupling proteins without transmembrane domains. A construct for the coexpression of *traD* and *yaa* has already been generated with a pCOLA-Duet vector. To simplify the purification another vector with a higher copy number and a tag on traD should be used.

To study the mechanism of DNA translocation in the neisserial T4SS it is crucial to understand the enzymatic activity and function of the key player TraD. To be able to draw conclusions about the mechanism of TraD in context with other characterized coupling proteins it is important to gain information about the oligomeric state of TraD, as well as its ATP binding and hydrolysis activity. The oligomeric state of TraD could be either studied with size exclusion chromatography or with dynamic light scattering in consideration of a possible influence on the oligomerization of ATP. The ATP binding activity and a possible binding affinity could be measured with fluorescence spectroscopy and anisotropy measurements with the ATP analogues Mant-ATP or TNP-ATP. The ATP hydrolysis activity of TraD could be studies with commonly established ATP hydrolysis assays (e.g. enzyme coupled ATP hydrolysis assay or malachite green assay).

Furthermore, it is most likely that TraD binds to ssT-DNA to enable the translocation through the secretion complex. Therefore, it is worth to test a possible DNA binding activity of TraD with different single and double stranded DNA substrates. Additionally, it should be studied whether TraD interact with other T4SS proteins of *N. gonorrhoeae*.

The MOB_H relaxase TraI

This study revealed first insights into the enzymatic activity of this novel relaxase family and provides enlightening information about the DNA processing activity, as well as on the domain organization and the oligomerization of TraI.

To gain further information about the DNA binding and cleavage activity of TraI it is important to identify the active site within the relaxase domain. It is most likely that TraI uses like other relaxases a tyrosine or phenylalanine for its cleavage reaction. Two conserved tyrosines (Y^{93} and Y^{201}) of TraI have been reported to be important for the secretion of DNA. Both tyrosines have been mutated to phenylalanine and purification trials have been made. Unfortunately, it was not possible up to now to isolate mutated TraI without aggregation. Nevertheless, the isolation of these two mutants should be in the focus of future experiments since there is strong evidence that these residues possibly resemble the active site of TraI. However, it is possible that both tyrosines do not participate in the catalytic site of TraI and an up to now unnoticed tyrosine or phenylalanine catalyzes the cleavage reaction. Initial evidences strengthening this assumption have been obtained in a structural comparison of a modeled structure of TraI and the structure of a relaxase like protein of *C. burnetii*. Comparing the active site of both proteins within a structural overlay, it is most likely that Y^{212} of TraI resembles the active site. Therefore, Y^{212} should be mutated to phenylalanine and the isolated protein should be tested for its DNA processing activity.

Information about the DNA binding affinity of TraI can be studied by determining different K_d values obtained in anisotropy measurements with different DNA substrates. Initial anisotropy measurements were already performed and indicate a high affinity of TraI on ssDNA. To study possible determinants for the sequence specificity of the cleavage reaction of TraI will be a challenging task for future research. Additional relaxosome components that could trigger the specificity of the cleavage reaction of TraI have not been reported for the neisserial T4SS. Nevertheless, it would be interesting to study a possible influence of the genome encoded IHF on the specificity of the cleavage reaction of TraI. Therefore, the IHF of *N. gonorrhoeae* should be recombinantly overproduced and purififed and should be added to the *in vitro* cleavage assays. Since up to now only the *oriT* regions of *N. gonorrhoeae* and from *E. coli* F-plasmid were tested for cleavage by TraI, the established cleavage assay should be performed using substrates containing *oriT* regions from other T4SSs like e.g. *A. tumefaciens*. Besides different DNA substrates and additional possible relaxosome components, the reaction conditions of the cleavage assay could be altered to obtain a specific reaction of TraI.

Furthermore, it is important to study the function of the distinct domains of TraI. Limited proteolysis revealed that TraI consist of at least three trypsin stable domains. Sequence alignments also indicate the presence of a conserved middle part in TraI_2 relaxases with a C-terminal DUF domain. To study the function of distinct domains of TraI different truncation of the protein have been generated and initial purification trails were performed. The truncated domains of TraI should be further studied considering DNA binding and processing, ATP hydrolysis activity, and possible protein-protein interactions. These constructs can be further used to characterize the function of the different domains in detail, e.g. the conserved HD sequence motifs in the relaxase domain can be mutated to study a possible influence on the metal binding mechanism. A possible metal dependent phosphatase activity of these sequence motifs can be studied e.g. with phosphatase assays offering different substrates. The truncated middle domain, as well as the relaxase domain can be tested for a

putative ATP dependent helicase activity by using the approach of molecular beacons (314). This approach was already used in initially experiments to study the helicase activity of the full-length protein. The putative function of the C-terminal DUF1528 domain in facilitating protein-protein interaction could be tested with various approaches like Surface Plasmon Resonance, different pull-down assays, cross-linking assays or size exclusion chromatography. Different possible binding partners like Yaf, TraC, Yaa or TraD could be tested for interaction with the C-terminal domain of TraI.

Besides the functional characterization of the different domains of TraI, the experimental approach with truncated TraI domains can be used to study the ssDNA transfer through the secretion complex in *N. gonorrhoeae*. Therefore, different mutant strains of *N. gonorrhoeae* carrying different truncations of *traI* could be generated and studied with regard to their levels of DNA secretion. This approach could give information about a putative C-terminal signal sequence important for the transfer of the processed DNA to the secretion complex.

To elucidate the DNA processing mechanism of TraI as a prototype of the uncharacterized MOB_H relaxases, it is important to gain besides biochemical data structural information about the protein. Therefore, crystallization attempts with the native full-length protein have been started in collaboration with Dr S. Smits from the group of Prof. Dr. L. Schmitt from the Heinrich Heine University in Düsseldorf.

The hypothetical protein Yaf

The biochemical characterization of Yaf revealed a nuclease activity that still needs to be studied further to understand its function within the context of DNA secretion in *N. gonorrhoeae*.

To obtain insights into the enzymatic function of Yaf, it is of outstanding importance to identify the catalytic residue and the metal binding site of Yaf. Previously performed mutagenesis screenings forced with the problem of low protein yields and the difficulty to choose an important residue. Based on recently obtained structural data of Yaf, it should be possible to identify the catalytic residue of the protein and to generate a non-functional mutant. The generated mutant can be used for a further biochemical characterization of the nuclease activity of Yaf.

The observed DNA processing activity of Yaf needs to be studied further to understand the still unknown function of this hypothetical protein. It is important to gain detailed information about the dsDNA degradation mechanism of Yaf (e.g. the minimal required dsDNA substrate length, a possible recognition site, and the determinants for specific degradation of linearized dsDNA, but not circularized DNA). Since weak evidence of dsDNA binding is given from the observed results of DNA binding and processing assay, further DNA binding experiments with changed reaction conditions and different DNA substrates should be performed.

A possible regulatory function of Yaf on the excision and integration of the GGI, similar to RDF proteins, can be studied in vivo with a *yaf* knock-out strain and the wild type strain. Both strains could be grown in certain stress conditions, most likely oxidative stress since *N. gonorrhoeae* is highly sensible to changing oxidative conditions and Yaf is a Mn^{2+} dependent protein, and the GGI excision frequency could be measured with quantitative PCR. A similar approach could be also used for the study of a possible regulatory influence on the expression level of the *dtr* genes (*traI*, *traD*, and also *yaa*).

7. Literature

1. Schoen C, Joseph B, Claus H, Vogel U, & Frosch M (2007) *Int J Med Microbiol* **297,** 601-613.
2. Grossman M, Sussman S, Gottfried D, Quock C, & Ticknor W (1964) *Am J Dis Child* **107,** 356-362.
3. Ward ME & Watt PJ (1972) *J Infect Dis* **126,** 601-605.
4. Maiden MC (2008) *Curr Opin Microbiol* **11,** 467-471.
5. Cohen MS (1998) *Lancet* **351 Suppl 3,** 5-7.
6. Zenilman JM (1998) *Curr Opin Infect Dis* **11,** 47-52.
7. Donovan B, Bodsworth NJ, Rohrsheim R, McNulty A, & Tapsall JW (2000) *Lancet* **355,** 1908.
8. Tapsall JW (2005) *Clin Infect Dis* **41 Suppl 4,** S263-268.
9. Biswas GD, Sox T, Blackman E, & Sparling PF (1977) *J Bacteriol* **129,** 983-992.
10. Hamilton HL & Dillard JP (2006) *Mol Microbiol* **59,** 376-385.
11. Davidsen T, Rodland EA, Lagesen K, Seeberg E, Rognes T, & Tonjum T (2004) *Nucleic Acids Res* **32,** 1050-1058.
12. Treangen TJ, Ambur OH, Tonjum T, & Rocha EP (2008) *Genome Biol* **9,** R60.
13. Davidsen T & Tonjum T (2006) *Nat Rev Microbiol* **4,** 11-22.
14. Dorward DW, Garon CF, & Judd RC (1989) *J Bacteriol* **171,** 2499-2505.
15. Wetzler LM, Blake MS, Barry K, & Gotschlich EC (1992) *J Infect Dis* **166,** 551-555.
16. Pettit RK & Judd RC (1992) *Mol Microbiol* **6,** 723-728.
17. Britigan BE, Cohen MS, & Sparling PF (1985) *N Engl J Med* **312,** 1683-1694.
18. Kahn ME, Barany F, & Smith HO (1983) *Proc Natl Acad Sci U S A* **80,** 6927-6931.
19. Barany F, Kahn ME, & Smith HO (1983) *Proc Natl Acad Sci U S A* **80,** 7274-7278.
20. Ohnishi M, Golparian D, Shimuta K, Saika T, Hoshina S, Iwasaku K, Nakayama S, Kitawaki J, & Unemo M (2011) *Antimicrob Agents Chemother* **55,** 3538-3545.
21. Ohnishi M, Saika T, Hoshina S, Iwasaku K, Nakayama S, Watanabe H, & Kitawaki J (2011) *Emerg Infect Dis* **17,** 148-149.
22. Tseng TT, Tyler BM, & Setubal JC (2009) *BMC Microbiol* **9 Suppl 1,** S2.
23. Ishii S, Kosaka T, Hori K, Hotta Y, & Watanabe K (2005) *Appl Environ Microbiol* **71,** 7838-7845.
24. Bardy SL, Ng SY, & Jarrell KF (2003) *Microbiology* **149,** 295-304.
25. Kato J, Kim HE, Takiguchi N, Kuroda A, & Ohtake H (2008) *J Biosci Bioeng* **106,** 1-7.
26. Christie PJ (2001) *Mol Microbiol* **40,** 294-305.
27. Shimoyama T, Kato S, Ishii S, & Watanabe K (2009) *Science* **323,** 1574.
28. Joshi H, Dave R, & Venugopalan VP (2009) *PLoS One* **4,** e6065.
29. Llosa M, Gomis-Ruth FX, Coll M, & de la Cruz Fd F (2002) *Mol Microbiol* **45,** 1-8.
30. Lawley TD, Klimke WA, Gubbins MJ, & Frost LS (2003) *FEMS Microbiol Lett* **224,** 1-15.
31. Hayes CS, Aoki SK, & Low DA (2010) *Annu Rev Genet* **44,** 71-90.
32. Abdallah AM, Gey van Pittius NC, Champion PA, Cox J, Luirink J, Vandenbroucke-Grauls CM, Appelmelk BJ, & Bitter W (2007) *Nat Rev Microbiol* **5,** 883-891.
33. Bleves S, Viarre V, Salacha R, Michel GP, Filloux A, & Voulhoux R (2010) *Int J Med Microbiol* **300,** 534-543.
34. Rego AT, Chandran V, & Waksman G (2010) *Biochem J* **425,** 475-488.
35. Papanikou E, Karamanou S, & Economou A (2007) *Nat Rev Microbiol* **5,** 839-851.
36. Driessen AJ & Nouwen N (2008) *Annu Rev Biochem* **77,** 643-667.
37. Hartl FU, Lecker S, Schiebel E, Hendrick JP, & Wickner W (1990) *Cell* **63,** 269-279.
38. Lycklama ANJA & Driessen AJ (2012) *Philos Trans R Soc Lond B Biol Sci* **367,** 1016-1028.
39. Muller M (2005) *Res Microbiol* **156,** 131-136.
40. Hutcheon GW & Bolhuis A (2003) *Biochem Soc Trans* **31,** 686-689.
41. Robinson C, Matos CF, Beck D, Ren C, Lawrence J, Vasisht N, & Mendel S (2011) *Biochim Biophys Acta* **1808,** 876-884.
42. Coulthurst SJ, Dawson A, Hunter WN, & Sargent F (2012) *Biochemistry* **51,** 1678-1686.
43. Delepelaire P (2004) *Biochim Biophys Acta* **1694,** 149-161.

44. Harley KT, Djordjevic GM, Tseng TT, & Saier MH (2000) *Mol Microbiol* **36,** 516-517.
45. Uhlen P, Laestadius A, Jahnukainen T, Soderblom T, Backhed F, Celsi G, Brismar H, Normark S, Aperia A, & Richter-Dahlfors A (2000) *Nature* **405,** 694-697.
46. Russo DM, Williams A, Edwards A, Posadas DM, Finnie C, Dankert M, Downie JA, & Zorreguieta A (2006) *Journal of Bacteriology* **188,** 4474-4486.
47. Duong F, Lazdunski A, & Murgier M (1996) *Mol Microbiol* **21,** 459-470.
48. Holland IB, Schmitt L, & Young J (2005) *Mol Membr Biol* **22,** 29-39.
49. Binet R, Letoffe S, Ghigo JM, Delepelaire P, & Wandersman C (1997) *Folia Microbiol (Praha)* **42,** 179-183.
50. Duong F, Bonnet E, Geli V, Lazdunski A, Murgier M, & Filloux A (2001) *Gene* **262,** 147-153.
51. Wandersman C & Delepelaire P (2004) *Annu Rev Microbiol* **58,** 611-647.
52. da Silva FG, Shen Y, Dardick C, Burdman S, Yadav RC, de Leon AL, & Ronald PC (2004) *Mol Plant Microbe Interact* **17,** 593-601.
53. Preston GM, Studholme DJ, & Caldelari I (2005) *FEMS Microbiol Rev* **29,** 331-360.
54. Cianciotto NP (2005) *Trends Microbiol* **13,** 581-588.
55. Sauvonnet N, Vignon G, Pugsley AP, & Gounon P (2000) *EMBO J* **19,** 2221-2228.
56. Kohler R, Schafer K, Muller S, Vignon G, Diederichs K, Philippsen A, Ringler P, Pugsley AP, Engel A, & Welte W (2004) *Mol Microbiol* **54,** 647-664.
57. Filloux A (2004) *Biochim Biophys Acta* **1694,** 163-179.
58. Ruer S, Stender S, Filloux A, & de Bentzmann S (2007) *J Bacteriol* **189,** 3547-3555.
59. Ball G, Durand E, Lazdunski A, & Filloux A (2002) *Mol Microbiol* **43,** 475-485.
60. Cornelis GR (2006) *Nat Rev Microbiol* **4,** 811-825.
61. Galan JE & Wolf-Watz H (2006) *Nature* **444,** 567-573.
62. Gophna U, Ron EZ, & Graur D (2003) *Gene* **312,** 151-163.
63. Mota LJ, Journet L, Sorg I, Agrain C, & Cornelis GR (2005) *Science* **307,** 1278.
64. Sekiya K, Ohishi M, Ogino T, Tamano K, Sasakawa C, & Abe A (2001) *Proc Natl Acad Sci U S A* **98,** 11638-11643.
65. Paul K, Erhardt M, Hirano T, Blair DF, & Hughes KT (2008) *Nature* **451,** 489-492.
66. Wilharm G, Lehmann V, Krauss K, Lehnert B, Richter S, Ruckdeschel K, Heesemann J, & Trulzsch K (2004) *Infect Immun* **72,** 4004-4009.
67. Veenendaal AK, Hodgkinson JL, Schwarzer L, Stabat D, Zenk SF, & Blocker AJ (2007) *Mol Microbiol* **63,** 1719-1730.
68. Akeda Y & Galan JE (2005) *Nature* **437,** 911-915.
69. Jacob-Dubuisson F, Fernandez R, & Coutte L (2004) *Biochim Biophys Acta* **1694,** 235-257.
70. Fan E, Fiedler S, Jacob-Dubuisson F, & Muller M (2012) *J Biol Chem* **287,** 2591-2599.
71. Newman CL & Stathopoulos C (2004) *Crit Rev Microbiol* **30,** 275-286.
72. Neil RB & Apicella MA (2009) *Infect Immun* **77,** 2285-2293.
73. Aoki SK, Poole SJ, Hayes CS, & Low DA (2011) *Virulence* **2,** 356-359.
74. Dautin N & Bernstein HD (2007) *Annu Rev Microbiol* **61,** 89-112.
75. Henderson IR, Navarro-Garcia F, Desvaux M, Fernandez RC, & Ala'Aldeen D (2004) *Microbiol Mol Biol Rev* **68,** 692-744.
76. Jacob-Dubuisson F, Locht C, & Antoine R (2001) *Mol Microbiol* **40,** 306-313.
77. Hodak H & Jacob-Dubuisson F (2007) *Res Microbiol* **158,** 631-637.
78. Clantin B, Hodak H, Willery E, Locht C, Jacob-Dubuisson F, & Villeret V (2004) *Proc Natl Acad Sci U S A* **101,** 6194-6199.
79. Aoki SK, Webb JS, Braaten BA, & Low DA (2009) *J Bacteriol* **191,** 1777-1786.
80. Nikolakakis K, Amber S, Wilbur JS, Diner EJ, Aoki SK, Poole SJ, Tuanyok A, Keim PS, Peacock S, Hayes CS*, et al.* (2012) *Mol Microbiol* **84,** 516-529.
81. Parsons HK, Vitovski S, & Sayers JR (2004) *Biochem Soc Trans* **32,** 1130-1132.
82. Leiman PG, Basler M, Ramagopal UA, Bonanno JB, Sauder JM, Pukatzki S, Burley SK, Almo SC, & Mekalanos JJ (2009) *Proc Natl Acad Sci U S A* **106,** 4154-4159.

83. Hood RD, Singh P, Hsu F, Guvener T, Carl MA, Trinidad RR, Silverman JM, Ohlson BB, Hicks KG, Plemel RL*, et al.* (2010) *Cell Host Microbe* **7,** 25-37.
84. Gibbs KA, Urbanowski ML, & Greenberg EP (2008) *Science* **321,** 256-259.
85. Wu HY, Chung PC, Shih HW, Wen SR, & Lai EM (2008) *J Bacteriol* **190,** 2841-2850.
86. Filloux A, Hachani A, & Bleves S (2008) *Microbiology* **154,** 1570-1583.
87. Records AR & Gross DC (2010) *J Bacteriol* **192,** 3584-3596.
88. Coros A, Callahan B, Battaglioli E, & Derbyshire KM (2008) *Mol Microbiol* **69,** 794-808.
89. Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, Honore N, Garnier T, Churcher C, Harris D*, et al.* (2001) *Nature* **409,** 1007-1011.
90. Smith J, Manoranjan J, Pan M, Bohsali A, Xu J, Liu J, McDonald KL, Szyk A, LaRonde-LeBlanc N, & Gao LY (2008) *Infect Immun* **76,** 5478-5487.
91. Converse SE & Cox JS (2005) *J Bacteriol* **187,** 1238-1245.
92. Burts ML, Williams WA, DeBord K, & Missiakas DM (2005) *Proc Natl Acad Sci U S A* **102,** 1169-1174.
93. Garufi G, Butler E, & Missiakas D (2008) *J Bacteriol* **190,** 7004-7011.
94. Bitter W, Houben EN, Bottai D, Brodin P, Brown EJ, Cox JS, Derbyshire K, Fortune SM, Gao LY, Liu J*, et al.* (2009) *PLoS Pathog* **5,** e1000507.
95. Gey Van Pittius NC, Gamieldien J, Hide W, Brown GD, Siezen RJ, & Beyers AD (2001) *Genome Biol* **2,** RESEARCH0044.
96. Pallen MJ (2002) *Trends Microbiol* **10,** 209-212.
97. Sao-Jose C, Baptista C, & Santos MA (2004) *J Bacteriol* **186,** 8337-8346.
98. Renshaw PS, Lightbody KL, Veverka V, Muskett FW, Kelly G, Frenkiel TA, Gordon SV, Hewinson RG, Burke B, Norman J*, et al.* (2005) *EMBO J* **24,** 2491-2498.
99. Renshaw PS, Panagiotidou P, Whelan A, Gordon SV, Hewinson RG, Williamson RA, & Carr MD (2002) *J Biol Chem* **277,** 21598-21603.
100. Champion PA, Stanley SA, Champion MM, Brown EJ, & Cox JS (2006) *Science* **313,** 1632-1636.
101. Ize B & Palmer T (2006) *Science* **313,** 1583-1584.
102. Christie PJ, Atmakuri K, Krishnamoorthy V, Jakubowski S, & Cascales E (2005) *Annu Rev Microbiol* **59,** 451-485.
103. Vergunst AC, van Lier MC, den Dulk-Ras A, Stuve TA, Ouwehand A, & Hooykaas PJ (2005) *Proc Natl Acad Sci U S A* **102,** 832-837.
104. Cascales E & Christie PJ (2003) *Nat Rev Microbiol* **1,** 137-149.
105. Alvarez-Martinez CE & Christie PJ (2009) *Microbiol Mol Biol Rev* **73,** 775-808.
106. de Paz HD, Sangari FJ, Bolland S, Garcia-Lobo JM, Dehio C, de la Cruz F, & Llosa M (2005) *Microbiology* **151,** 3505-3516.
107. McCullen CA & Binns AN (2006) *Annu Rev Cell Dev Biol* **22,** 101-127.
108. Williams JJ & Hergenrother PJ (2008) *Curr Opin Chem Biol* **12,** 389-399.
109. Novick RP (1987) *Microbiol Rev* **51,** 381-395.
110. Garcillan-Barcia MP, Francia MV, & de la Cruz F (2009) *FEMS Microbiol Rev* **33,** 657-687.
111. Guglielmini J, Quintais L, Garcillan-Barcia MP, de la Cruz F, & Rocha EP (2011) *PLoS Genet* **7,** e1002222.
112. Guglielmini J, de la Cruz F, & Rocha EP (2012) *Mol Biol Evol*.
113. Dillard JP & Seifert HS (2001) *Mol Microbiol* **41,** 263-277.
114. Burrus V, Pavlovic G, Decaris B, & Guedon G (2002) *Mol Microbiol* **46,** 601-610.
115. Burrus V & Waldor MK (2004) *Res Microbiol* **155,** 376-386.
116. Christie PJ (2004) *Biochim Biophys Acta* **1694,** 219-234.
117. Pansegrau W & Lanka E (1996) *Prog Nucleic Acid Res Mol Biol* **54,** 197-251.
118. Schroder G & Lanka E (2005) *Plasmid* **54,** 1-25.
119. Chandran V, Fronzes R, Duquerroy S, Cronin N, Navaza J, & Waksman G (2009) *Nature* **462,** 1011-1015.
120. Fronzes R, Schafer E, Wang L, Saibil HR, Orlova EV, & Waksman G (2009) *Science* **323,** 266-268.
121. Burns DL (2003) *Curr Opin Microbiol* **6,** 29-34.

122. Shrivastava R & Miller JF (2009) *Curr Opin Microbiol* **12,** 88-93.
123. Terradot L & Waksman G (2011) *FEBS J* **278,** 1213-1222.
124. Celli J & Gorvel JP (2004) *Curr Opin Microbiol* **7,** 93-97.
125. Bacon DJ, Alm RA, Burr DH, Hu L, Kopecko DJ, Ewing CP, Trust TJ, & Guerry P (2000) *Infect Immun* **68,** 4384-4390.
126. Hofreuter D, Odenbreit S, & Haas R (2001) *Mol Microbiol* **41,** 379-391.
127. Clarke M, Maddera L, Harris RL, & Silverman PM (2008) *Proc Natl Acad Sci U S A* **105,** 17978-17981.
128. Middleton R, Sjolander K, Krishnamurthy N, Foley J, & Zambryski P (2005) *Proc Natl Acad Sci U S A* **102,** 1685-1690.
129. Wallden K, Williams R, Yan J, Lian PW, Wang L, Thalassinos K, Orlova EV, & Waksman G (2012) *Proc Natl Acad Sci U S A* **109,** 11348-11353.
130. Krause S, Pansegrau W, Lurz R, de la Cruz F, & Lanka E (2000) *J Bacteriol* **182,** 2761-2770.
131. Rashkova S, Spudich GM, & Christie PJ (1997) *J Bacteriol* **179,** 583-591.
132. Yeo HJ, Savvides SN, Herr AB, Lanka E, & Waksman G (2000) *Mol Cell* **6,** 1461-1472.
133. Yuan Q, Carle A, Gao C, Sivanesan D, Aly KA, Hoppner C, Krall L, Domke N, & Baron C (2005) *J Biol Chem* **280,** 26349-26359.
134. Berger BR & Christie PJ (1994) *J Bacteriol* **176,** 3646-3660.
135. Judd PK, Kumar RB, & Das A (2005) *Mol Microbiol* **55,** 115-124.
136. Jakubowski SJ, Krishnamoorthy V, & Christie PJ (2003) *J Bacteriol* **185,** 2867-2878.
137. Cascales E & Christie PJ (2004) *Science* **304,** 1170-1173.
138. Audette GF, Manchak J, Beatty P, Klimke WA, & Frost LS (2007) *Microbiology* **153,** 442-451.
139. Judd PK, Kumar RB, & Das A (2005) *Proc Natl Acad Sci U S A* **102,** 11498-11503.
140. Kumar RB, Xie YH, & Das A (2000) *Mol Microbiol* **36,** 608-617.
141. Jakubowski SJ, Kerr JE, Garza I, Krishnamoorthy V, Bayliss R, Waksman G, & Christie PJ (2009) *Mol Microbiol* **71,** 779-794.
142. Cascales E & Christie PJ (2004) *Proc Natl Acad Sci U S A* **101,** 17228-17233.
143. Zahrl D, Wagner M, Bischof K, Bayer M, Zavecz B, Beranek A, Ruckenstuhl C, Zarfel GE, & Koraimann G (2005) *Microbiology* **151,** 3455-3467.
144. Anderson LB, Hertzel AV, & Das A (1996) *Proc Natl Acad Sci U S A* **93,** 8889-8894.
145. Jakubowski SJ, Cascales E, Krishnamoorthy V, & Christie PJ (2005) *J Bacteriol* **187,** 3486-3495.
146. Christie PJ & Cascales E (2005) *Mol Membr Biol* **22,** 51-61.
147. van Ulsen P & Tommassen J (2006) *FEMS Microbiol Rev* **30,** 292-319.
148. Horstman AL & Kuehn MJ (2000) *J Biol Chem* **275,** 12489-12496.
149. Thompson SA, Wang LL, West A, & Sparling PF (1993) *J Bacteriol* **175,** 811-818.
150. Woodhams KL, Benet ZL, Blonsky SE, Hackett KT, & Dillard JP (2012) *J Bacteriol* **194,** 2275-2285.
151. Pohlner J, Halter R, Beyreuther K, & Meyer TF (1987) *Nature* **325,** 458-462.
152. Hauck CR & Meyer TF (1997) *FEBS Lett* **405,** 86-90.
153. Senior BW, Stewart WW, Galloway C, & Kerr MA (2001) *J Infect Dis* **184,** 922-925.
154. Hadi HA, Wooldridge KG, Robinson K, & Ala'Aldeen DA (2001) *Mol Microbiol* **41,** 611-623.
155. Serruto D, Adu-Bobie J, Scarselli M, Veggi D, Pizza M, Rappuoli R, & Arico B (2003) *Mol Microbiol* **48,** 323-334.
156. van Ulsen P, van Alphen L, ten Hove J, Fransen F, van der Ley P, & Tommassen J (2003) *Mol Microbiol* **50,** 1017-1030.
157. Brigulla M & Wackernagel W (2010) *Appl Microbiol Biotechnol* **86,** 1027-1041.
158. Brochet M, Rusniok C, Couve E, Dramsi S, Poyart C, Trieu-Cuot P, Kunst F, & Glaser P (2008) *Proc Natl Acad Sci U S A* **105,** 15961-15966.
159. Wozniak RA & Waldor MK (2010) *Nat Rev Microbiol* **8,** 552-563.
160. Franke AE & Clewell DB (1981) *Cold Spring Harb Symp Quant Biol* **45 Pt 1,** 77-80.
161. Shoemaker NB, Barber RD, & Salyers AA (1989) *J Bacteriol* **171,** 1294-1302.
162. Hochhut B & Waldor MK (1999) *Mol Microbiol* **32,** 99-110.

163. Douard G, Praud K, Cloeckaert A, & Doublet B (2010) *PLoS One* **5,** e15302.
164. Daccord A, Ceccarelli D, & Burrus V (2010) *Mol Microbiol* **78,** 576-588.
165. van der Meer JR & Sentchilo V (2003) *Curr Opin Biotechnol* **14,** 248-254.
166. Waldor MK, Tschape H, & Mekalanos JJ (1996) *J Bacteriol* **178,** 4157-4165.
167. Davies MR, Shera J, Van Domselaar GH, Sriprakash KS, & McMillan DJ (2009) *J Bacteriol* **191,** 2257-2265.
168. Ravatn R, Studer S, Springael D, Zehnder AJ, & van der Meer JR (1998) *J Bacteriol* **180,** 4360-4369.
169. Sullivan JT & Ronson CW (1998) *Proc Natl Acad Sci U S A* **95,** 5145-5149.
170. He J, Baldini RL, Deziel E, Saucier M, Zhang Q, Liberati NT, Lee D, Urbach J, Goodman HM, & Rahme LG (2004) *Proc Natl Acad Sci U S A* **101,** 2530-2535.
171. Drenkard E & Ausubel FM (2002) *Nature* **416,** 740-743.
172. Kikuchi Y & Nash HA (1979) *Proc Natl Acad Sci U S A* **76,** 3760-3764.
173. Ravatn R, Studer S, Zehnder AJ, & van der Meer JR (1998) *J Bacteriol* **180,** 5505-5514.
174. Ramsay JP, Sullivan JT, Stuart GS, Lamont IL, & Ronson CW (2006) *Mol Microbiol* **62,** 723-734.
175. Mullany P, Wilks M, Lamb I, Clayton C, Wren B, & Tabaqchali S (1990) *J Gen Microbiol* **136,** 1343-1349.
176. Haren L, Ton-Hoang B, & Chandler M (1999) *Annu Rev Microbiol* **53,** 245-281.
177. Qiu X, Gurkar AU, & Lory S (2006) *Proc Natl Acad Sci U S A* **103,** 19830-19835.
178. Mohd-Zain Z, Turner SL, Cerdeno-Tarraga AM, Lilley AK, Inzana TJ, Duncan AJ, Harding RM, Hood DW, Peto TE, & Crook DW (2004) *J Bacteriol* **186,** 8114-8122.
179. Lewis JA & Hatfull GF (2001) *Nucleic Acids Res* **29,** 2205-2216.
180. Sutanto Y, Shoemaker NB, Gardner JF, & Salyers AA (2002) *Mol Microbiol* **46,** 1239-1246.
181. Li M, Shen X, Yan J, Han H, Zheng B, Liu D, Cheng H, Zhao Y, Rao X, Wang C*, et al.* (2011) *Mol Microbiol* **79,** 1670-1683.
182. Lee PS & Grossman AD (2006) *Mol Microbiol* **60,** 853-869.
183. Wozniak RA & Waldor MK (2009) *PLoS Genet* **5,** e1000439.
184. Ceccarelli D, Daccord A, Rene M, & Burrus V (2008) *J Bacteriol* **190,** 5328-5338.
185. Vogelmann J, Ammelburg M, Finger C, Guezguez J, Linke D, Flotenmeyer M, Stierhof YD, Wohlleben W, & Muth G (2011) *EMBO J* **30,** 2246-2254.
186. Marrero J & Waldor MK (2007) *J Bacteriol* **189,** 3302-3305.
187. Auchtung JM, Lee CA, Garrison KL, & Grossman AD (2007) *Mol Microbiol* **64,** 1515-1528.
188. Mullany P, Roberts AP, & Wang H (2002) *Cell Mol Life Sci* **59,** 2017-2022.
189. Dominguez NM, Hackett KT, & Dillard JP (2011) *J Bacteriol* **193,** 377-388.
190. Piazzolla D, Cali S, Spoldi E, Forti F, Sala C, Magnoni F, Deho G, & Ghisotti D (2006) *J Gen Virol* **87,** 2423-2431.
191. Niwa O, Yamagishi H, & Ozeki H (1978) *Mol Gen Genet* **159,** 259-268.
192. Sentchilo V, Czechowska K, Pradervand N, Minoia M, Miyazaki R, & van der Meer JR (2009) *Mol Microbiol* **72,** 1293-1306.
193. Perrin A, Nassif X, & Tinsley C (1999) *Infect Immun* **67,** 6119-6129.
194. Snyder LA, Jarvis SA, & Saunders NJ (2005) *Microbiology* **151,** 4005-4013.
195. McShan WM, Williams RP, & Hull RA (1987) *Infect Immun* **55,** 3017-3022.
196. Zola TA, Strange HR, Dominguez NM, Dillard JP, & Cornelissen CN (2010) *Infect Immun* **78,** 2429-2437.
197. Karlin S (1998) *Curr Opin Microbiol* **1,** 598-610.
198. Hamilton HL, Dominguez NM, Schwartz KJ, Hackett KT, & Dillard JP (2005) *Mol Microbiol* **55,** 1704-1721.
199. Smith HO, Gwinn ML, & Salzberg SL (1999) *Res Microbiol* **150,** 603-616.
200. Ramsey ME, Woodhams KL, & Dillard JP (2011) *Front Microbiol* **2,** 61.
201. Kroll JS, Wilks KE, Farrant JL, & Langford PR (1998) *Proc Natl Acad Sci U S A* **95,** 12381-12385.
202. Salgado-Pabon W, Jain S, Turner N, van der Does C, & Dillard JP (2007) *Mol Microbiol* **66,** 930-947.

203. Jain S, Zweig M, Peeters E, Siewering K, Hackett KT, Dillard JP, & van der Does C (2012) *PLoS One* **7,** e35285.
204. Krieg AM (1999) *Biochim Biophys Acta* **1489,** 107-116.
205. Hamilton HL, Schwartz KJ, & Dillard JP (2001) *J Bacteriol* **183,** 4718-4726.
206. Frost LS, Ippen-Ihler K, & Skurray RA (1994) *Microbiol Rev* **58,** 162-210.
207. Wozniak RA, Fouts DE, Spagnoletti M, Colombo MM, Ceccarelli D, Garriss G, Dery C, Burrus V, & Waldor MK (2009) *PLoS Genet* **5,** e1000786.
208. Salgado-Pabon W, Du Y, Hackett KT, Lyons KM, Arvidson CG, & Dillard JP (2010) *J Bacteriol* **192,** 1912-1920.
209. Pansegrau W, Balzer D, Kruft V, Lurz R, & Lanka E (1990) *Proc Natl Acad Sci U S A* **87,** 6555-6559.
210. Howard MT, Nelson WC, & Matson SW (1995) *J Biol Chem* **270,** 28381-28386.
211. Guogas LM, Kennedy SA, Lee JH, & Redinbo MR (2009) *J Mol Biol* **386,** 554-568.
212. Byrd DR & Matson SW (1997) *Mol Microbiol* **25,** 1011-1022.
213. Pansegrau W, Ziegelin G, & Lanka E (1990) *J Biol Chem* **265,** 10637-10644.
214. Pansegrau W & Lanka E (1996) *J Biol Chem* **271,** 13068-13076.
215. Yoshida H, Furuya N, Lin YJ, Guntert P, Komano T, & Kainosho M (2008) *J Mol Biol* **384,** 690-701.
216. Ziegelin G, Furste JP, & Lanka E (1989) *J Biol Chem* **264,** 11989-11994.
217. Nelson WC, Howard MT, Sherman JA, & Matson SW (1995) *J Biol Chem* **270,** 28374-28380.
218. Ragonese H, Haisch D, Villareal E, Choi JH, & Matson SW (2007) *Mol Microbiol* **63,** 1173-1184.
219. Moncalian G, Grandoso G, Llosa M, & de la Cruz F (1997) *J Mol Biol* **270,** 188-200.
220. Lu J, Wong JJ, Edwards RA, Manchak J, Frost LS, & Glover JN (2008) *Mol Microbiol* **70,** 89-99.
221. Ziegelin G, Pansegrau W, Lurz R, & Lanka E (1992) *J Biol Chem* **267,** 17279-17286.
222. Toro N, Datta A, Carmi OA, Young C, Prusti RK, & Nester EW (1989) *J Bacteriol* **171,** 6845-6849.
223. Fekete RA & Frost LS (2002) *J Biol Chem* **277,** 16705-16711.
224. Williams SL & Schildbach JF (2007) *J Bacteriol* **189,** 3813-3823.
225. Williams SL & Schildbach JF (2006) *Nucleic Acids Res* **34,** 426-435.
226. Paterson ES & Iyer VN (1992) *J Bacteriol* **174,** 499-507.
227. Guasch A, Lucas M, Moncalian G, Cabezas M, Perez-Luque R, Gomis-Ruth FX, de la Cruz F, & Coll M (2003) *Nat Struct Biol* **10,** 1002-1010.
228. Travers A (1997) *Curr Biol* **7,** R252-254.
229. Sut MV, Mihajlovic S, Lang S, Gruber CJ, & Zechner EL (2009) *J Bacteriol* **191,** 6888-6899.
230. Gomis-Ruth FX, Sola M, de la Cruz F, & Coll M (2004) *Curr Pharm Des* **10,** 1551-1565.
231. Lanka E & Wilkins BM (1995) *Annu Rev Biochem* **64,** 141-169.
232. Datta S, Larkin C, & Schildbach JF (2003) *Structure* **11,** 1369-1379.
233. Lang S, Gruber K, Mihajlovic S, Arnold R, Gruber CJ, Steinlechner S, Jehl MA, Rattei T, Frohlich KU, & Zechner EL (2010) *Mol Microbiol* **78,** 1539-1555.
234. Francia MV, Varsaki A, Garcillan-Barcia MP, Latorre A, Drainas C, & de la Cruz F (2004) *FEMS Microbiol Rev* **28,** 79-100.
235. Lahue EE & Matson SW (1988) *J Biol Chem* **263,** 3208-3215.
236. Scherzinger E, Lurz R, Otto S, & Dobrinski B (1992) *Nucleic Acids Res* **20,** 41-48.
237. Pansegrau W, Schroder W, & Lanka E (1993) *Proc Natl Acad Sci U S A* **90,** 2925-2929.
238. Matson SW, Nelson WC, & Morton BS (1993) *J Bacteriol* **175,** 2599-2606.
239. Grandoso G, Llosa M, Zabala JC, & de la Cruz F (1994) *Eur J Biochem* **226,** 403-412.
240. Fukuda H & Ohtsubo E (1995) *J Biol Chem* **270,** 21319-21325.
241. Nunez B & De La Cruz F (2001) *Mol Microbiol* **39,** 1088-1099.
242. Pansegrau W & Lanka E (1991) *Nucleic Acids Res* **19,** 3455.
243. Pansegrau W, Schroder W, & Lanka E (1994) *J Biol Chem* **269,** 2782-2789.
244. Stern JC & Schildbach JF (2001) *Biochemistry* **40,** 11586-11595.
245. Monzingo AF, Ozburn A, Xia S, Meyer RJ, & Robertus JD (2007) *J Mol Biol* **366,** 165-178.

246. Larkin C, Datta S, Harley MJ, Anderson BJ, Ebie A, Hargreaves V, & Schildbach JF (2005) *Structure* **13,** 1533-1544.
247. Dyda F & Hickman AB (2003) *Structure* **11,** 1310-1311.
248. Ilyina TV & Koonin EV (1992) *Nucleic Acids Res* **20,** 3279-3285.
249. Zimmerman MD, Proudfoot M, Yakunin A, & Minor W (2008) *J Mol Biol* **378,** 215-226.
250. Varsaki A, Moncalian G, Garcillan-Barcia Mdel P, Drainas C, & de la Cruz F (2009) *J Bacteriol* **191,** 1446-1455.
251. Close TJ, Tait RC, Rempel HC, Hirooka T, Kim L, & Kado CI (1987) *J Bacteriol* **169,** 2336-2344.
252. Nelson WC, Morton BS, Lahue EE, & Matson SW (1993) *J Bacteriol* **175,** 2221-2228.
253. Silverman PM & Sholl A (1996) *J Bacteriol* **178,** 5787-5789.
254. Chen Y, Zhang X, Manias D, Yeo HJ, Dunny GM, & Christie PJ (2008) *J Bacteriol* **190,** 3632-3645.
255. Atmakuri K, Cascales E, Burton OT, Banta LM, & Christie PJ (2007) *EMBO J* **26,** 2540-2551.
256. Disque-Kochem C & Dreiseikelmann B (1997) *J Bacteriol* **179,** 6133-6137.
257. Zhang S & Meyer R (1997) *Mol Microbiol* **25,** 509-516.
258. Schildbach JF, Robinson CR, & Sauer RT (1998) *J Biol Chem* **273,** 1329-1333.
259. Moncalian G & de la Cruz F (2004) *Biochim Biophys Acta* **1701,** 15-23.
260. Lu J, den Dulk-Ras A, Hooykaas PJ, & Glover JN (2009) *Proc Natl Acad Sci U S A* **106,** 9643-9648.
261. Tsai MM, Fu YH, & Deonier RC (1990) *J Bacteriol* **172,** 4603-4609.
262. Di Laurenzio L, Frost LS, & Paranchych W (1992) *Mol Microbiol* **6,** 2951-2959.
263. Bowie JU & Sauer RT (1990) *J Mol Biol* **211,** 5-6.
264. Gomis-Ruth FX, Sola M, Acebo P, Parraga A, Guasch A, Eritja R, Gonzalez A, Espinosa M, del Solar G, & Coll M (1998) *EMBO J* **17,** 7404-7415.
265. Schreiter ER & Drennan CL (2007) *Nat Rev Microbiol* **5,** 710-720.
266. Beranek A, Zettl M, Lorenzoni K, Schauer A, Manhart M, & Koraimann G (2004) *J Bacteriol* **186,** 6999-7006.
267. Locht C, Coutte L, & Mielcarek N (2011) *FEBS J* **278,** 4668-4682.
268. Cabezon E, Sastre JI, & de la Cruz F (1997) *Mol Gen Genet* **254,** 400-406.
269. Gomis-Ruth FX, Moncalian G, de la Cruz F, & Coll M (2002) *J Biol Chem* **277,** 7556-7566.
270. Gomis-Ruth FX & Coll M (2001) *Int J Biochem Cell Biol* **33,** 839-843.
271. Kumar RB & Das A (2002) *Mol Microbiol* **43,** 1523-1532.
272. Moncalian G, Cabezon E, Alkorta I, Valle M, Moro F, Valpuesta JM, Goni FM, & de La Cruz F (1999) *J Biol Chem* **274,** 36117-36124.
273. Balzer D, Pansegrau W, & Lanka E (1994) *J Bacteriol* **176,** 4285-4295.
274. Gomis-Ruth FX, Moncalian G, Perez-Luque R, Gonzalez A, Cabezon E, de la Cruz F, & Coll M (2001) *Nature* **409,** 637-641.
275. Lee MH, Kosuk N, Bailey J, Traxler B, & Manoil C (1999) *J Bacteriol* **181,** 6108-6113.
276. Hormaeche I, Iloro I, Arrondo JL, Goni FM, de la Cruz F, & Alkorta I (2004) *J Biol Chem* **279,** 10955-10961.
277. Tato I, Matilla I, Arechaga I, Zunzunegui S, de la Cruz F, & Cabezon E (2007) *J Biol Chem* **282,** 25569-25576.
278. Tato I, Zunzunegui S, de la Cruz F, & Cabezon E (2005) *Proc Natl Acad Sci U S A* **102,** 8156-8161.
279. Abajy MY, Kopec J, Schiwon K, Burzynski M, Doring M, Bohn C, & Grohmann E (2007) *J Bacteriol* **189,** 2487-2496.
280. Hamilton CM, Lee H, Li PL, Cook DM, Piper KR, von Bodman SB, Lanka E, Ream W, & Farrand SK (2000) *J Bacteriol* **182,** 1541-1548.
281. Studier FW (2005) *Protein Expr Purif* **41,** 207-234.
282. Taylor RG, Walker DC, & McInnes RR (1993) *Nucleic Acids Res* **21,** 1677-1678.
283. Studier FW & Moffatt BA (1986) *J Mol Biol* **189,** 113-130.
284. Wild J & Szybalski W (2004) *Methods Mol Biol* **267,** 155-167.
285. Miroux B & Walker JE (1996) *J Mol Biol* **260,** 289-298.

286. Ferrer M, Chernikova TN, Yakimov MM, Golyshin PN, & Timmis KN (2003) *Nat Biotechnol* **21,** 1266-1267.
287. Hartinger D, Heinl S, Schwartz HE, Grabherr R, Schatzmayr G, Haltrich D, & Moll WD (2010) *Microb Cell Fact* **9,** 62.
288. Giacalone MJ, Gentile AM, Lovitt BT, Berkley NL, Gunderson CW, & Surber MW (2006) *Biotechniques* **40,** 355-364.
289. Grunberg-Manago M (1999) *Annu Rev Genet* **33,** 193-227.
290. Kido M, Yamanaka K, Mitani T, Niki H, Ogura T, & Hiraga S (1996) *J Bacteriol* **178,** 3917-3925.
291. Waugh DS (2011) *Protein Expr Purif* **80,** 283-293.
292. Nesterenko MV, Tilley M, & Upton SJ (1994) *J Biochem Biophys Methods* **28,** 239-242.
293. Domanico PL & Tse-Dinh YC (1991) *J Inorg Biochem* **42,** 87-96.
294. Zhu CX, Roche CJ, & Tse-Dinh YC (1997) *J Biol Chem* **272,** 16206-16210.
295. VanAken T, Foxall-VanAken S, Castleman S, & Ferguson-Miller S (1986) *Methods Enzymol* **125,** 27-35.
296. Hekman K, Guja K, Larkin C, & Schildbach JF (2008) *Nucleic Acids Res* **36,** 4565-4572.
297. Cheton PL & Archibald FS (1988) *Free Radic Biol Med* **5,** 325-333.
298. Coassin M, Ursini F, & Bindoli A (1992) *Arch Biochem Biophys* **299,** 330-333.
299. Kopec J, Bergmann A, Fritz G, Grohmann E, & Keller W (2005) *Biochem J* **387,** 401-409.
300. Street LM, Harley MJ, Stern JC, Larkin C, Williams SL, Miller DL, Dohm JA, Rodgers ME, & Schildbach JF (2003) *Biochim Biophys Acta* **1646,** 86-99.
301. Nash RP, Habibi S, Cheng Y, Lujan SA, & Redinbo MR (2010) *Nucleic Acids Res* **38,** 5929-5943.
302. Xu YC & Bremer H (1997) *Nucleic Acids Res* **25,** 4067-4071.
303. Matson SW & Ragonese H (2005) *J Bacteriol* **187,** 697-706.
304. Lucas M, Gonzalez-Perez B, Cabezas M, Moncalian G, Rivas G, & de la Cruz F (2010) *J Biol Chem* **285,** 8918-8926.
305. Ranquet C, Ollagnier-de-Choudens S, Loiseau L, Barras F, & Fontecave M (2007) *J Biol Chem* **282,** 30442-30451.
306. Wu HJ, Seib KL, Srikhanta YN, Edwards J, Kidd SP, Maguire TL, Hamilton A, Pan KT, Hsiao HH, Yao CW*, et al.* (2010) *J Proteomics* **73,** 899-916.
307. Papp-Wallace KM & Maguire ME (2006) *Annu Rev Microbiol* **60,** 187-209.
308. Michalke B, Halbach S, & Nischwitz V (2007) *J Environ Monit* **9,** 650-656.
309. Archibald FS & Duong MN (1984) *J Bacteriol* **158,** 1-8.
310. Medicis ED, Paquette J, Gauthier JJ, & Shapcott D (1986) *Appl Environ Microbiol* **52,** 567-573.
311. Leviatan S, Sawada K, Moriyama Y, & Nelson N (2010) *J Biol Chem* **285,** 23548-23556.
312. McCoy J & La Ville E (2001) *Curr Protoc Protein Sci* **Chapter 6,** Unit 6 7.
313. Drew D, Lerch M, Kunji E, Slotboom DJ, & de Gier JW (2006) *Nat Methods* **3,** 303-313.
314. Belon CA & Frick DN (2008) *Biotechniques* **45,** 433-440, 442.

Index of Tables

Index of Figures

Abbreviations

APS	Ammonium persulfate
BN-PAGE	blue native polyacrylamide gel electrophoresis
CV	column volume
Chaps	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
DDM	n-dodecyl-α-D-maltopyranoside
DM	n-decyl-α-D-maltopyranoside
DTT	Dithiothreitol
dsDNA	double stranded DNA
ssDNA	sing
frac.	frac
GI	genetic island
Hfr	high frequency of recombination
ICE	integrative conjugative element
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KPP	Potassium phosphate buffer
LDAO	Lauryldimethylamine-oxide
MBP	maltose binding protein
MOB class	mobilization class
MPF	maiting pair formation
MALDI-TOF	matrix-assisted laser desorption/ionization – time of flight analysis
Nano-LC	nano – liquid chromatography
oriT	origin of transfer region
PNK	Polynucleotide kinase
PVDF	Polyvinylidene fluoride
PCR	polymerase chain reaction
PMSF	Phenylmethylsulfonyl fluoride
RDF	recombination directionality factors
SBA	Sodium boric acid buffer + EDTA
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEM	scanning electron microscopy
SOC	super optimal broth medium
TB	Tris boric acid buffer
TCA	Trichloroacetic acid
TGS	Tris-Glycine-SDS buffer
TEMED	Tetramethylethylenediamine
TEV	tobacco etch virus
TxSS	type X secretion system
TxS	type X secretion
UDM	Undecyl-β-D-maltoside
WT	wild type